

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 28, 2002, 12:10:01 ; Search time 3887.7 Seconds  
(without alignments)  
11758.542 Million cell updates/sec

Title: US-09-674-824-1  
Perfect score: 2771  
Sequence: 1 cgcacatccactgccttgc.....aaaaaaaaaaaaaaaaaaaaa 2771

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues  
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.om.\*
- 20: em.or.\*
- 21: em.ov.\*
- 22: em.pat.\*
- 23: em.ph.\*
- 24: em.pl.\*
- 25: em.ro.\*
- 26: em.sts.\*
- 27: em.sy.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htgo.hum.\*
- 31: em.htgo.inv.\*
- 32: em.htgo.rod.\*
- 33: em.htg.hum.\*
- 34: em.htg.inv.\*
- 35: em.htg.rod.\*
- 36: em.htg.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2771	100.0	2771	6	AX010492
2	2541.8	91.7	2662	6	AX031272 Sequence
3	2519	90.9	2591	8	AF091803
4	2486.6	89.7	2575	8	AF091803 Triticum
5	2332.8	84.2	2421	8	TAE292521 Triticum
6	2239	80.8	2239	8	TAE292522 Triticum
7	2054	74.1	2055	8	A93350 Sequence 1
8	1393	50.3	2383	8	TAU48227 Triticum ae
9	1366	49.3	2391	8	A93359 Sequence 1
10	1364.4	49.2	2592	8	AF036891 Zea mays
11	1332	48.1	2590	8	AF168786 Sorghum b
12	1329.4	48.0	2533	6	AR049919 Sequence
13	1329.4	48.0	2533	6	E06904 cDNA encodi
14	1329.4	48.0	2533	6	DI6202 Rice mRNA f
15	1327.6	47.9	2592	8	AF165890 Oryza sat
16	1325.6	47.8	2592	8	A63308 Sequence 1
17	1248.4	45.1	1620	6	AR106494 Sequence
18	752.8	27.2	2360	6	AR106498 Sequence
19	752.8	27.2	2360	6	AR112888 Sequence 9
20	752.8	27.2	2360	6	AR112888 Sequence
21	648.4	23.4	1758	6	STSTASINT
22	648.4	23.4	1758	6	A51201 Sequence 3
23	596	21.5	11387	8	AR112885 Sequence
24	567.6	20.5	10336	6	AF091802 Aegilops
25	431	15.6	8387	8	AX031274 Sequence
26	349	12.6	2380	6	AF234163 Sequence
27	349	12.6	2380	6	AR049921 Sequence
28	299	10.8	5072	6	A63310 Sequence 3
29	287	10.4	2793	6	AX031276 Sequence
30	287	10.4	2793	6	A51205 Sequence 7
31	284.4	10.3	2325	8	AR112887 Sequence
32	283.2	10.2	2988	8	AF068834 Ipomoea b
33	279	10.1	1926	6	PSSTASINT
34	279	10.1	1926	6	A51203 Sequence 5
35	275.6	9.9	2903	8	AR112886 Sequence
36	264.2	9.5	8900	8	AF383878 Oryza sat
c 37	264.2	9.5	170371	8	RICE1
38	255.8	9.2	2595	8	AB026295 Oryza sat
c 39	253.8	9.2	256191	2	STSTARSYN
40	252.6	9.1	93735	8	AC084046 Trypanoso
41	249.4	9.0	2575	8	ATAC008261 Arabidops
42	248.2	9.0	2394	8	AF173900 Manihot e
43	233	8.4	2097	6	AF395537 Oryza sat
44	233	8.4	2480	8	AR106493 Sequence
45	232	8.4	2865	8	AF019297 Zea mays
					AF019296 Zea mays

ALIGNMENTS

RESULT 1	AX010492	2771 bp	DNA	PAT	06-SEP-2000
LOCUS	Sequence 1 from Patent WO9958688.				
DEFINITION	AX010492				
ACCESSION	AX010492.1	GI:9997335			
VERSION					
KEYWORDS	bread wheat.				
SOURCE	Triticum aestivum				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.				
REFERENCE	1 (bases 1 to 2771)				
AUTHORS	Luetticke,S., Block,M. and Loerz,H.				
TITLE	Nucleic acid molecules which code for enzymes derived from wheat and which are involved in the synthesis of starch				
JOURNAL	Patent: WO 9958688-A 1 18-NOV-1999;				
FEATURES	LUETTICKE STEPHANIE (DE); BLOCK MARTINA (DE); LOERZ HORST (DE); HOECHST SCHERING AGREVO GMBH (DE) Location/Qualifiers				



```

source      1. .2771
            /organism="Triticum aestivum"
            /db_xref="taxon:4565"
CDS         280..2550
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAC07516.1"
            /translation="WNAITGAGAGCLASVRLRADPATAAASACVVRAARLRLRLLARGRY
            VAEISREGFAAPAAQQAQQQLAPPVPGFLAPEPPAPAQPASPATQPPLPDAGVGELADPL
            LKGTIAEDSIDSIIVAAEQSEIDANEQEQAKVTRISIVFTGERAKYPAKSGLGDV
            CGSLPIALAAGRHVVMVMPYLRNGSSDKNIKALYTAKHIKPCFGSSHEVTFFHFEY
            RDNVDVPHDPSYHRPQSJLYGDNFAGFDNQFRYTLLCYAAEAPILLELGQIYYGP
            NMNFVNDHMSILVPVLAAYKYPGYVRDSRSTLVINHLAHQCVEPASTYPDGLQP
            EMTGALWFVPEWARHALDGEAVNFKLGAVTADRTVTVSQGVSWETVAEGGGOL
            NEUUSRRKSYLINGINGIDINDWNPTDKCLPHHYSDDDLSGKAACKAELOKEGLPV
            REDVPLIGFIRLDYQKDIDLTKMAIPELMRDEQVFVNLGSDPIIFGHWRSSTESSKY
            DKTRGMVGSVPVSHRTIATGCDLILLMRSFPBCPNQLNYAQMYGTIVPVHGTGGLRDT
            VETFNPFKAGEBGTGWAFSPDLTYDKMLWALTAMSTFREHKHSWEGLMKRMGTMDRT
            WDHAESPSTRSSGPSWTNPSTSCRGLGRSKCESPSALKTSSSSFRGPEGYPTLCRP
            ATVESOCACILFAGESTYTDGCAAAATVASGRQLQFMGIRKCAAGWLTAKHHSIDGS
            LSVRVTAIEIRNQLVLT"
BASE COUNT   669 a    699 c    754 g    649 t
ORIGIN
Query Match       100.0%; Score 2771; DB 6; Length 2771;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2771; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1  Cgcaactcactcgcttggccccactcccactctctctctccgcgcacacagtcgagcgc 60
Db 1  CGCCACTCCACTCGCTTTGCCCACTCCCAGTCPTCTCCCCCGCACACCAGAGTCGCG 60

Qy 61 acggctctaccatcacctcgcgctcggccaccggcaaacccccgatccgcgttttg 120
Db 61 ACCGGCTCATCAACCATCACTTGCCCTCGGCACCGCGCAAACCCCGCATCGCTTTTG 120

Qy 121 caggcagcgcaactaaacccgggagcgccgcccgcgagcagcagcagcagtcgag 180
Db 121 CAGCGAGCGCACTAAAAACCCGGGAGCGCGCGCGGCGGAGCAGCAGCACCAGCTGGG 180

Qy 181 agagagagccttcgccccggccgcacagcagggggagatccaccgtccgctcgctcgca 240
Db 181 AGAGAGAGCGCTTCGCCCCCGCCGCCACCGAGCGGGGGCGATCCACCCTGCTGGTCCGCA 240

Qy 241 cctctcgcctcctcctctcgtccgcgcgcgccacacccattggcggcagcggcgctcgcc 300
Db 241 CCTCTTCGGCTCTCTCCCTGTCCCGCGCGGCCACACCATGGGCGGCGAGCGGGCTCGGC 300

Qy 301 gccgggtgcctcgcccccgagtcgcgctcgcgcgcatcggcgagcagcgcgccggggcg 360
Db 301 GCCGGGTGCTCGCCCCCACAAGCTTCGCGCTTCGCGCGCGATCCGGCGAGCGGGGGCGCG 360

Qy 361 tcgcgctgcgctctcgcgcgcgcctcggcgcttggcgcgggggccgctacctgcgcgag 420
Db 361 TCGCGCTGCGCTCTGTCGCGCGCGGCTTCGCGCTTGGCGCGGGGGCGCGCTACGTGCGCGAG 420

Qy 421 ctacgagggagggccccgcggcgccccgcgcgagcagcagcaactggcccccgccctc 480
Db 421 CTCAGCAGGGAGGGCCCCCGGGCGCGCCCCCGCGCAGCAGCAGCAACTGGCCCCCGCGCTC 480

Qy 481 gtgcagagcttcctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcag 540
Db 481 GTGCAGAGCTTCTTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 540

Qy 541 ccgcccctgcgcgagcgcgcgctgggggaactcgcgcgcgcgcctcctcgtcctaaggagatt 600
Db 541 CCGCCCCCTGCGGAGCGCGCGGCTGGGGAACTCGCGCGCGGACCTCCTCTGCTCGAAGGATT 600

Qy 601 gctgagattccatcgacagcataattgtgctgcgaagtgcagagattcttcagatcatg 660
Db 601 GCTGAGATTCCATCGACAGCAATAATGTGGCTGCAAGTGACAGGATTCGTGAATCATG 660

```



QY 1741 atgcttgatcggggatcccaatttttaagcctggatgagatctaccgagtcgagttac 1800  
 Db 1741 ATGCTTGGATCGGGATCCCAATTTTGAAGCTGGATGAGATACCGAGTCGAGTTAC 1800  
 QY 1801 aaggaataattccgtggtggtttagtttccagtttccacagaaataactgca 1860  
 Db 1801 AAGGATAAATTCGGTGGATGGTGGATTTAGTGTTCAGTTTCCACAGAAATCTGCA 1860  
 QY 1861 gttgcatatattgtaataccatcagatttgaaccttgctgttataatcagctatat 1920  
 Db 1861 GTTGCGGATATATGTTAATGCCATCAGATTTGAACCTTGGCTTAAATACGCTATAT 1920  
 QY 1921 gctatgaatagttacagtttctcttagtttccagtttgggctcgcagacacagtc 1980  
 Db 1921 CTTATGCAATATGTATAGTTCCTGTAGTTTCATGGAACCTGGGGCTTCGACACAGTC 1980  
 QY 1981 gagacctcaaccttttgggtgcaaaagagagagaggttacagggtgggtcttcacgc 2040  
 Db 1981 GAGACCTTCAACCTTTTGGTGCAAAGAGAGAGAGGGTACAGGTTGGCGTCTCAACCG 2040  
 QY 2041 ctaacgtggaagatgttgggtcattgcgaacccgcatgtcgacattcagggagcac 2100  
 Db 2041 CTAACCGTGACAAAGATGTTGGGCATTCGAACCGCATGTCCACATTCAGGAGCAC 2100  
 QY 2101 aagccgtccctgggagggctcatgaagcagcagcagcagcagcagcagcagcagcagc 2160  
 Db 2101 AAGCCGTCTCGGAGGGCTCATGAAGCAGGCGATGACGAAGACATACATGCGGACCAT 2160  
 QY 2161 gccccagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2220  
 Db 2161 GCCCCAGCAGTACGAGCAGATCTTCGAGTGGCGCTTCGTGGACCAACCTTACGTCATGT 2220  
 QY 2221 agaggggagctgggaggtcccaagtcgaggtctcttcaagtcctgaagacatctcttca 2280  
 Db 2221 AGAGGGGAGCTGGGAGGTCCAAAGTCCGAGTCTCTTCAGCTCTGAAGACATCTCTTCA 2280  
 QY 2281 tcttccgagcggccggaaggtacccctgtacattggttctctctctacagtagatcg 2340  
 Db 2281 TCTTCCCGCGCCGGAAGGATACCCCTGTACATTGGTGTCTGTCTACAGTAGAGTCG 2340  
 QY 2341 caatgcgctgtcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 2400  
 Db 2341 CAATCGCCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400  
 QY 2401 gcggtgacagctcgggtgagcagcttacagttttggggaataaggaaggtatgct 2460  
 Db 2401 GCGGTGACAGCTTCGGGTGGAGCAGCTTACAGTTTGGGAATATAGGAAGGATGTGCT 2460  
 QY 2461 gcaggtggttaacagcaaacaccactcagatggcagcctctcttctcgtgttacagct 2520  
 Db 2461 GCAGGATGTTTACAGCAAAAGCACCCTCAGATGGCAGCCTCTCTGCTCGTGTACAGCT 2520  
 QY 2521 gaaatcgaacaactcgttgaactctttagcctttagcttagcttagcttagcttagcttag 2580  
 Db 2521 GAAATCAGAAACCAACTGTTGACTCTTAGCTTAGCTTAGCTTAGCTTAGCTTAGCTTAGCT 2580  
 QY 2581 gtgtatgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 2640  
 Db 2581 GTGTATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2640  
 QY 2641 ctgctgtttttttttaaatacaagaggggtttcttctcagatttcaataaaaaa 2700  
 Db 2641 CTGCTGTTTTTTTTTAAATCAAAAGAGGGGTTTCTCTCCGATTTCAATTAATAAAAAA 2700  
 QY 2701 aaaaaa 2771  
 Db 2701 AAAAAA 2771  
 QY 2761 aaaaaa 2771  
 Db 2761 AAAAAA 2771

RESULT 2  
 AX031272  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 AX031272  
 Sequence 11 from Patent WO9914314.  
 AX031272  
 AX031272.1 GI:10278604  
 Aegilops tauschii.  
 Aegilops tauschii  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Pooideae; Triticeae; Aegilops.  
 1 (bases 1 to 2662)  
 Li, Z., Morell, M. and Rahman, S.  
 Regulation of gene expression in plants  
 Patent: WO 9914314-A 11 25-MAR-1999;  
 GOODMAN FIELDER LTD (AU); LI ZHONGYI (AU); MORELL MATTHEW (AU);  
 RAHMAN SADQUR (AU); UNIV AUSTRALIAN (AU); COMMW SCIENT IND RES  
 ORG (AU); GROUPE LINAGRAIN PACIFIC PTY L (AU)  
 FEATURES  
 source  
 1..2662  
 /organism="Aegilops tauschii"  
 /db\_xref="taxon:37682"  
 /issue\_type="ENDOSPERM"  
 misc\_feature  
 1..2651  
 /product="NUCLEOTIDE SEQUENCE OF CDNA WHEAT SSS I"  
 BASE COUNT 592 a 677 c 754 g 639 t  
 ORIGIN  
 Query Match 91.7%; Score 2541.8; DB 6; Length 2662;  
 Best Local Similarity 98.6%; Pred. No. 0;  
 Matches 2597; Conservative 0; Mismatches 32; Indels 6; Gaps 3;  
 QY 25 ctccacactctctctcccgccgacacacagtcgagccggtctatcaccatcacctcg 84  
 Db 2 CTCCACACTCTCTCTCCCGCGCACACCGAGTCGGCACCGGCTCATCACCATCACCTCG 61  
 QY 85 gctcggccacggcgaacaccccccgatctcttctgagcagcagcactaaacccggg 144  
 Db 62 GCCTGGCCACCGGCAACACCCCGGATCCGCTTTTGGCAGCAGCGCACTAAACCCGGG 121  
 QY 145 gagcgcccccgccgagcagcagcagcagtcgggagagagaggttcgccccgccccg 204  
 Db 122 GAGCGCGCCCGCGCGAGCAGCAGCAGCAGTGGGAGAGAGAGGTTGCCGCCGCCGCCG 181  
 QY 205 caccgagcgggcgatccacgtctcgtgctccgacactctcctcctcctcctcctcctc 264  
 Db 182 CACGAGCGGGGCGATCCACGTCGGTGGCTCGCGACCTCTCTCCGCTCTCTCCCTGTCC 241  
 QY 265 cgcgccccacacccatggcgagcgggcgctcgcgccccgggtgctcgcacagcgctc 324  
 Db 242 CGCGCGCCACACCCATGGCGGCGAGCGGCGTGGCGCGCGGTTGCTCGCCGCCGCCG 301  
 QY 325 cgctgcgccccgacccgagcgcgcccgggcgcccgccgctcgtcgtcgtcgtcgtcgtc 384  
 Db 302 CGCTCGCGCGCGATCCGCGAGCGCGCGCGCGGCTCGCGCTCGCGCTCGCGCGCGG 361  
 QY 385 ctccgcgctctgcgcgcgccgctacgtcgcgagctcagcagggagggccccgccccg 444  
 Db 362 CTCGCGGCTTGGCGGGGGCGCGCTACGTTGCCGAGCTAGCAGGAGGCGGCCGCCG 421  
 QY 445 cgccccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 504  
 Db 422 CGCCCCCGCAGCAGCAGCAACTGGCCCGCGCTCTGTCGACAGGCTTCTCGCGCGCG 481  
 QY 505 ccgccccgcccccgagtcgccccgccccgccccgccccgccccgccccgccccgccccg 564  
 Db 482 CCG 541  
 QY 565 ggggaactcgccccgacctctcgtcgaaggtatgctgaggtattcattcagacagcata 624  
 Db 542 GGGGAACTCG 601



Qy	1705	attccagagctcatgaggagacgctgcaatttggatgcttggatcggggtcgaatt	1764
Db	1682	attccagagctcatgaggagacgctgcaatttggatgcttggatcggggtcgaatt	1741
Qy	1765	tttgaagcttgagatgagatcctaccgagtcgagttacaagagataaaatccgtgga	1824
Db	1742	tttgaagcttgagatgagatcctaccgagtcgagttacaagagataaaatccgtgga	1801
Qy	1825	ggatttagtctccagtttcccccacagaaatactgcaggggttgagatatattgta	1884
Db	1802	ggatttagtctccagtttcccccacagaaatactgcaggggttgagatatattgta	1861
Qy	1885	tcgagattggaacctcgcgtcttaatacagctatgctatgcaataatggtacagtt	1944
Db	1862	tcgagattggaacctcgcgtcttaatacagctatgctatgcaataatggtacagtt	1921
Qy	1945	gtagttcatggaacctcgcgtcttaatacagctatgctatgcaataatggtacagtt	2004
Db	1922	gtagttcatggaacctcgcgtcttaatacagctatgctatgcaataatggtacagtt	1981
Qy	2005	aaagagagaggggttaacaggggttcacaggggttcacaggggttcacaggggttc	2064
Db	1982	aaagagagaggggttaacaggggttcacaggggttcacaggggttcacaggggttc	2041
Qy	2065	gcatttcgaacccgcgctgctgacatttcagggagcacaagccgctcctggggggtc	2124
Db	2042	gcatttcgaacccgcgctgctgacatttcagggagcacaagccgctcctggggggtc	2101
Qy	2125	aaagagagaggggttaacaggggttcacaggggttcacaggggttcacaggggttc	2183
Db	2102	aaagagagaggggttaacaggggttcacaggggttcacaggggttcacaggggttc	2161
Qy	2184	ttcgaatggggttcgctgacacccctacgctcagtagagcggggaactggggggtc	2243
Db	2162	ttcgaatggggttcgctgacacccctacgctcagtagagcggggaactggggggtc	2221
Qy	2244	gtcgagctccttcagctcctgagacacccctcctcctcctcctcctcctcctcct	2303
Db	2222	gtcgagctccttcagctcctgagacacccctcctcctcctcctcctcctcctcct	2281
Qy	2304	ccctgtacattcgtctcctcctcctcctcctcctcctcctcctcctcctcctcct	2363
Db	2282	ccctgtacattcgtctcctcctcctcctcctcctcctcctcctcctcctcctcct	2340
Qy	2364	gcggttcgagacacatgacgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	2423
Db	2341	gcggttcgagacacatgacgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	2400
Qy	2424	acagttacagtttggggaataaggaagggatgctgctgagggatgctgctgagga	2483
Db	2401	acagttacagtttggggaataaggaagggatgctgctgagggatgctgctgagga	2460
Qy	2484	ccactcagatggcagcctcctcgttccggttaccagctgaaatcagaaacacagtg	2543
Db	2461	ccactcagatggcagcctcctcgttccggttaccagctgaaatcagaaacacagtg	2520
Qy	2544	tccttagccttaggagttggaagtttgcctcctcctcctcctcctcctcctcct	2603
Db	2521	tccttagccttaggagttggaagtttgcctcctcctcctcctcctcctcctcct	2580
Qy	2604	tgacaaattattgacgctgctggagaaat----tttatcttgcctgctgctgctg	2654
Db	2581	tgacaaattattgacgctgctggagaaat----tttatcttgcctgctgctgctg	2635
RESULT	3		
LOCUS	AF091803	2591 bp mRNA	PLN 14-SEP-1999
DEFINITION	Trilicium aestivum starch synthase I mRNA, complete cds.		
ACCESSION	AF091803		
VERSION	AF091803.1	GI:5880465	
KEYWORDS	bread wheat.		
Db	1625	atttggtgcaagtgagcagagattctgagatcagtgagtgcaatgagcaacctcaagct	684
Qy	1602	atttggtgcaagtgagcagagattctgagatcagtgagtgcaatgagcaacctcaagct	661
Db	1685	aaagttacacgttagcatcgtgttggtagctgggtagagctgctccttatagcgaagtc	744
Qy	1662	aaagttacacgttagcatcgtgttggtagctgggtagagctgctccttatagcgaagtc	721
Db	1745	gggttgggaagatgttgggttggtagctgggtagagctgctccttatagcgaagtc	804
Qy	1722	gggttgggaagatgttgggttggtagctgggtagagctgctccttatagcgaagtc	781
Db	1805	atgatttgaatgcaagatacttaaatgggtcctctgataaaactatgcgaagggcatc	864
Qy	1782	atgatttgaatgcaagatacttaaatgggtcctctgataaaactatgcgaagggcatc	841
Db	1865	tacactggaagcaaatgaagttccatgcttgggggagatcacatgaagtgaccttttt	924
Qy	1842	tacactggaagcaaatgaagttccatgcttgggggagatcacatgaagtgaccttttt	901
Db	1925	catgagtatagacacacgctcgattgggtgttgcgacatcctcgttcatacagacca	984
Qy	1902	catgagtatagacacacgctcgattgggtgttgcgacatcctcgttcatacagacca	961
Db	1985	ggaagtttataggagataaattttgggtccttgggtgataatcagttcagatcacac	1044
Qy	1962	ggaagtttataggagataaattttgggtccttgggtgataatcagttcagatcacac	1021
Db	2045	ctttgctatgctcagcagagcccaactaatccttgggtgagagagatatttatgga	1104
Qy	2022	ctttgctatgctcagcagagcccaactaatccttgggtgagagagatatttatgga	1081
Db	2105	cagaattgcatgttgggtgaaacgattggcagccctgtgacagctcctcgtgct	1164
Qy	2082	cagaattgcatgttgggtgaaacgattggcagccctgtgacagctcctcgtgct	1141
Db	2165	gcaaaatagacacacaggttttacagagattcccgagcagccttgggttatcataat	1224
Qy	2142	gcaaaatagacacacaggttttacagagattcccgagcagccttgggttatcataat	1201
Db	2225	ttagcacatcaggtgtggagcctgcaagtacatacctcgtatcgtgggttgcctcctg	1284
Qy	2202	ttagcacatcaggtgtggagcctgcaagtacatacctcgtatcgtgggttgcctcctg	1261
Db	2285	tggtatggagcttttagaattgggtatttcagagtgagcagagggatgctccttgaca	1344
Qy	2262	tggtatggagcttttagaattgggtatttcagagtgagcagagggatgctccttgaca	1321
Db	2345	ggtgagggcagtttaacttttgaagggagcagttgtgacagcagatcggatttgaccgc	1404
Qy	2322	ggtgagggcagtttaacttttgaagggagcagttgtgacagcagatcggatttgaccgc	1381
Db	2405	agtcagggctttacatgggaggtcacaactgctgaaggtgagagggcctcctcaatgagct	1464
Qy	2382	agtcagggctttacatgggaggtcacaactgctgaaggtgagagggcctcctcaatgagct	1441
Db	2465	ttcaagctccgaaaaagtgattgaaatgggaattgaaatgggaattgcaattgagttgg	1524
Qy	2442	ttcaagctccgaaaaagtgattgaaatgggaattgaaatgggaattgcaattgagttgg	1501
Db	2525	aaccacacacagcaagtgctcctcctcctcctcctcctcctcctcctcctcctcct	1584
Qy	2502	aaccacacacagcaagtgctcctcctcctcctcctcctcctcctcctcctcctcct	1561
Db	2585	gcaaatgtaaaagctgaaattgagagaggttgggtttacacgttaagggagatgctcct	1644
Qy	2562	gcaaatgtaaaagctgaaattgagagaggttgggtttacacgttaagggagatgctcct	1621
Db	2645	ctgattggctttatggaagacgtgattaccagaaagcattgctcctcatttaaatggcc	1704
Qy	2622	ctgattggctttatggaagacgtgattaccagaaagcattgctcctcatttaaatggcc	1681



















Db	718	CGAAGCACATTAAGAAATCCATGCTTTGGGGGATCACATGAAGTGACCTTTTTCATGAG	777
Qy	931	tatagacacaacgtcgattgggtgtttgtcgatcatccgtacatcacagacacgaagt	990
Db	778	TATAGACACAACGTCGATTGGGTGTGTGTCGATCATCCGTGATATACAGACACGAGAAGT	837
Qy	991	ttatatggagataatttgggtgcttttggtygataatcagttcagatcacacactctttgc	1050
Db	838	TTATATGGAGATAAATTTTGGTGTCTTTGGTGATAATCAGTTACAGATACACACTCCTTTCG	897
Qy	1051	tatgctgcacgagagcccccactaatccttgaattggaggagataatttatggacagaat	1110
Db	898	TATGCTGCATGGAGAGCCCACTAATCCTTGAATGGGAGGATATATTATATGGACAAAT	957
Qy	1111	tgcatgttggtttgaacagatggcatggcagccttgtccagctcctttcttgcgtcaaaa	1170
Db	958	TGCATGTGTTGTGAACGATTTGCGATGCGCAGCTTGTGCCAGTCTCTTTCGTGCACAAA	1017
Qy	1171	tatagaccatacgggtgtttacagagattcccgacagacaccttggtatatacataattagca	1230
Db	1018	TATAGACCATACGGTGTACAGAGATTCGCCGACAGACCTTTGTTATACATAATTTAGCA	1077
Qy	1231	catcagggctggagcctgcaagtacatactctgactggagatgcctcctgaatggtat	1290
Db	1078	CATCAGGGTGTGAGCCTGCAAGTACATATCCTGATCTGGGATGCCCTTCCTGAATGGTAT	1137
Qy	1291	ggagctcttgaatgggtatttccagaataggccaaggagcatgcccttgacaagggtgag	1350
Db	1138	GGAGCTTTAGAAATGGGTATTTCAGAAATGGGCAAGAGGCGATGCCCTTGACAAGGTGAG	1197
Qy	1351	gcagttaacttttgaaggagcagttgtgacagcagatcgatgtgacccgtgacccgtcagtcag	1410
Db	1198	GCAGTTAACTTTTGAAGAGGAGCAGTTGTGACAGCAGATCGGATTTGTGACCGTCAGCTAG	1257
Qy	1411	ggttattcaatggaggctcaaacctgctgaagtgagacagggccctcaatgagctcttaagc	1470
Db	1258	GGTTATTCNTGGGAGGTCACAAGTGCCTGAAGGTGGACAGGCGCTCAATGAGCTCTTAAAGC	1317
Qy	1471	tccggaaaagtgatttgaatgaaattgcaatgaaattgacattaatgattggaaacccc	1530
Db	1318	TCCCGAAAAGTGTATTGAATGGAATTGTAAATGGAATTGACATTAATGATTGGAAACCC	1377
Qy	1531	accacagaacagtgctccctcatcattattctgcatgacctctctgtgaaaggccaaa	1590
Db	1378	ACCACAGCAAGTGTCTCCCTCATATTATTTCTGCGATGACCTCTCTGGAAGGCCAAA	1437
Qy	1591	tgtaaactgaaattgcagaaggagtgtgggtttaccttaactgaaggagagatgtcctctgatt	1650
Db	1438	TGTAAAGCTCAATTGCAGAAGAGTGGGTTTACTGTAAAGGAGGATGTTCTCTCTGATT	1497
Qy	1651	ggcttattggaagactgattaccagaagccaattgatctcaataaaatggccattcca	1710
Db	1498	GGCTTATTGTGAAGACTGGATTACCAGAAAGGCATTGATCTCATTAATAATGGCCATTCCA	1557
Qy	1711	gagctatgagggagacgtgcaatttgtcatctgtgattctggtgaggtcccaatttttga	1770
Db	1558	GAGCTCATGAGSGAGGACCTGCAATTTGTCTGCTGGATCTGGGGATCCAAATTTTGAA	1617
Qy	1771	ggctgatgatctaccgagctgagttacaagataaaatcccgtaggtgggttggattt	1830
Db	1618	GGCTGGATGAGATCTACCGAGTCCAGTTACAAGGATAAATCCGTGGATGGGTGGGATTT	1677
Qy	1831	agtgttccagtttcccacagataactgcaggttgcgataattgttaatgccatcaga	1890
Db	1678	AGTGTTCAGTTTCCCACAGAAATNACTGCAAGGTTCGATATATTGTTAATGCCATCGAGA	1737
Qy	1891	tttgaaccttgcggtctttaaactgactatatgtcatgcaatatggtacagtttccctgaatt	1950
Db	1738	TTTGAACCTTGGGTGCTTAATCAGCTATATGCTATGCAATATGGTACAGTTTCCCTGTAGTT	1797
Qy	1951	catggaactggggcctccagagacacagtcgagaccttaaccccttttgggtgcaaaagga	2010

[illegible]

CPs



/db\_xref="GI:6741618"  
/translation="TOPPLDPAAGVGEIAPDLLEGIAEDSIDSIIVASQDSEIMDA  
NQPOAKVTRSTVFVTEAAPYAKSGGLDVCGLPDLAARHVRVVMVPRYLNGSS  
DKNYAKALYAKHIKPCFGGSHVETFEHYRDNVDWVFDHPSYHRPGLYDNFGA  
PDGNFRYLICYACEAPLILELGGYIYGONCMFVNDWHASLVPVLAAKARYPGV  
YRDSRSTLIHNLHQGVPEASTPDGLGIPENYCALEWVPEWARHRLDKGAVNF  
LKGVAVTADRIYVTVSGVSWETTAEGGGGLNELLSRKSVLNGVINGIDINDWNPPT  
DKCLPHYSVDLSLQAKCAKAPLQELGLPREDVPLGIFGRIDYQKGLDILIKAP  
ELMRDVOPLMGLSGDPIFEQWMRSTESSYKDFRWGWFVSPVSHRITAGCDLILMP  
SRAPECLNOLYAMQYTVPVVHGTGRLDVTETFPNPFKAKEEGTGWAFSPLVDKM  
LVALRTMSTREHKPSWEGMLKGMKTWDHAPSTSRSSGSPSWNPSTCRRL  
GRSKCESPALYTSSSSPGPEGYPTLRCPATVESOCACLLWFAGSRTYDGCACAAV  
TAGGGHOLFWRIGKCAACWLTAKHSDSLSVRVTAIRNLQVL"  
BASE COUNT 611 a 448 c 590 g 590 t  
ORIGIN

Query Match		80.8%	Score 2239;	DB 6;	Length 2239;		
Best Local Similarity		100.0%	Pred. No. 5e-292;				
Matches 2239;		Conservative	0;	Mismatches	0;	Indels	0;
Qy	533	cgacgcagccgcccctgcgagccgctggggggaactcgcgcgcacccctcctctcg	592				
Db	1	CGAGCCAGCCGCCCTCGCGGACGCCGCGTGGGGAACTCGCGCCGACCTCCTGCTCG	60				
Qy	593	aagggattcgtgaggtatccatcacagcataatttgctgctcaagtgcagcagattctg	652				
Db	61	AAGGGATTGCTGAGGATTCCATCAGACGACATAATTTGGTGGCTGCAAGTGACGAGATTCTG	120				
Qy	653	agatcatgagtcgaatgagcaactcaagctcaagttacacgtagcatcgtgtttgtga	712				
Db	121	AGATCATGGATGCGAATGAGCAACCTCAAGCTAAAGTTACACGTAGCATCGTGTTCGTA	180				
Qy	713	ctggtgaagctcctccttatgcaaatgcaagggggttgagagattgttggttcgttac	772				
Db	181	CTGGTGAAGCTCCTCTTATGCAAAAGTCAGGGGGTGGGAGATGTTGGTTCGTAC	240				
Qy	773	caatgctcctgctcgtggtccaccgagtgatggttgtaattgcaagatacactaaatg	832				
Db	241	CAATTGCTCTTCTGCTGCTGCTACCGAGTGATGTTGTTAATGCCAAGATACTTAAATG	300				
Qy	833	ggtcctctgataaaactatgcaagggattatcacactgcgaagcacataagattccat	892				
Db	301	GGTCCCTCTGATAAAACTATGCAAAAGCATATACACTGCCAAGCACATTAAGATTCCAT	360				
Qy	893	gctttggggatcacatgaagtgaaccttttcatgagtagtagacaacgtcgattggg	952				
Db	361	GCTTTGGGGGATCACATGAAGTGACCTTTTTCATGAGTATAGACACAACGTCGATTGGG	420				
Qy	953	tgtttgctgatcatccgtcatatcacagaccaggagtttatatgagataatttgggtg	1012				
Db	421	TGTTTGTGATCATCCGTATATCAGACACAGGAAGTTTATATGAGATAATTTTGGTG	480				
Qy	1013	cttttgggtgataatcagttcagatacacactccttctgatatcgtcatcgagggcccccac	1072				
Db	481	CTTTTGGTGATAATCATAGTTTCAGATACACACTCCTTTGCTATGCTGATCGAGGCCCCNAC	540				
Qy	1073	taatcctggaattggagagataatttatgacagaattgcatgtttgttggaacgatt	1132				
Db	541	TAATTCCTTGAATTGGGAGGATATATTTATGACAGAAATTCATGTTGTTGTGAAACGATT	600				
Qy	1133	ggcatgcagccttgtgcccagtccttcttctgctgcaaaatatagaccatacgggtgtttaca	1192				
Db	601	GGCATGCAGCCCTTGTGCCAGTCTCTTCTGCTGCAAAATATAGACCATACGGGTGTTTACA	660				
Qy	1193	gagattcccgagcacaccctgttatatacaataatttagcacatcagggtgtgagccctgcaa	1252				
Db	661	GAGATTCCCGCAGCACCCCTTGTATTATCATATTTTAGCACATCAGGGTGGAGCCCTGCAA	720				
Qy	1253	gtacatactcgtgactggtgctcctcctgaatggtatgagattttagaatgggtatttc	1312				
Db	721	GTACATATCCCTGATCTGGGATTGCTCCTGAAATGGTATGAGCTTTAGAAATGGGTATTC	780				

Qy	1313	cagaatgggcaagagagcatgccccttgacaaggggtgagcagtttaactttttgaaagag	1372				
Db	781	CAGAAATGGCAAGAGGAGCATGCCCTTGACAAGGGTGAGGAGCTTAACCTTTTGAAGAGAG	840				
Qy	1373	cagttgacagcagatcgagattgtaacgctcagtcagggttattcatgggggtcacaa	1432				
Db	841	CAGTTGTGACAGCAGATCGGATTTGACCCCTCAGTCAGGGTTATTTCATGGAGGTCACAA	900				
Qy	1433	ctgctgaaggtggacaggggctcaatgagctcttaagctcccccacaaagtgtattgaatg	1492				
Db	901	CTGCTGAAGTGGACAGGGCTCAATGAGCTCTTAAGCTCCCGAAAGATGTATTGAATG	960				
Qy	1493	gaattgtaaatggaattgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat	1552				
Db	961	GAATTGTAATGGAATTCACATTAATGATTGGAACCCACACACAGACAGCAAGTGTCTCCCTC	1020				
Qy	1553	atcattattctgcatgacccctctctgaaaggccaaatgtaaaagtgaattgcagaag	1612				
Db	1021	ATCATATTCTGTCGATGACCTCTCTGAAAGGCCAAATGTAAAGCTGAAATTCAGAAAG	1080				
Qy	1613	agttgggtttaccctgtaaggagagatttctctctgattggctttatttggagactgatt	1672				
Db	1081	AGTTGGGTTTACCTGTAAGGGAGGATGTTCTCTGATTGGCTTTTATTGGAAAGACTGAT	1140				
Qy	1673	accgaaagagcattgaatctcattaaatggccattccagagctcatgagggagacgtgc	1732				
Db	1141	ACCAGAAAGGCATGATCTCATTAATGAGCCATTCAGAGCTCATGAGGGAGACGTGC	1200				
Qy	1733	aatttgcattgctgagctggggtcgggtcccaatttttgaaggctggagatgagactaccag	1792				
Db	1201	AATTTGTCATGCTGCTGATCTGGGATCCAAATTTTGAAGGCTGGATGAGATCTACCGAGT	1260				
Qy	1793	cgagttacaagataaaatccgtgagtggttgagattttagtctccagttcccccacagaa	1852				
Db	1261	CGAGTTACAAAGGATAAATTTCCGTGGATGGGTGGATTTAGTGTCCACTTTCCACAGAA	1320				
Qy	1853	taactcaggttgcgatatattgttaatgccatcgagattgaaacctgcgggtcttaate	1912				
Db	1321	TAATCTCAGGTTGCGATATATTGTTAATGCCATCGAGATTTTGAACCTTCCGGCTCTTAATC	1380				
Qy	1913	agctatgctatgcaatatgtaacattcctgtagttcatggaactgggggctcccgag	1972				
Db	1381	AGCTATATGCTATGCAATATGTTGATACAGTTTCTGTTAGTTCATGGAATGGGGGCTCCGAG	1440				
Qy	1973	acagtcagagacccctcaaccccttttggtaaaaggagaggggtacaggggtggggtc	2032				
Db	1441	ACAGTCTGAGACCTTCAACCCCTTTTGGTGCATAAGAGAGAGGAGGTACAGGGTGGGGCT	1500				
Qy	2033	tctcaccgctaaacgctggacaagatgttgggcatggaacccgagatgctcgacattca	2092				
Db	1501	TCTCACCGCTAACCGTGGACAAGATGTTGTCGATTTGGCAACCCGATGTCGACATTCAC	1560				
Qy	2093	gggagcaacagccgctcctgggaggggtcctgaagcgagggcatgacgaaagacatacgt	2152				
Db	1561	GGGAGCAACAAAGCCGCTCTGGGAGGGGCTCATGAAGCGAGGATGACGAAAGACATACGT	1620				
Qy	2153	gggacctggcccgagcagtcacgagcagatcttcgagtggggcttcgtggacacaccta	2212				
Db	1621	GGGACCATGCCCGCAGCAGTACGAGCAGATACCCCTGCTACATTGCGTGTCTCTACAG	1680				
Qy	2213	cgtcatgtagcggggactggggaggtcccaagtcgaagtcgaagtcctcctcagctcgaagacat	2272				
Db	1681	CGTCATGTAGAGCGGACATGGGGAGGTCCTCAAGTGGAGTCTCTCTTACAGTCTCTACAG	1740				
Qy	2273	ccttcctcctcctcctgggggggggggaggaataccctcgtatcattgcttgcgtacag	2332				
Db	1741	CCTCTTCAATCTTCCGCGGCCCGGAGGATACCCCTGCTACATTGCGTGTCTCTACAG	1800				
Qy	2333	tagatgcgaatgagcctcgtgttgggttcgcccgttcgagaaacataagcggctgtg	2392				
Db	1801	TAGAGTCGCAATGCGCCCTGCTTGTGTTGGTTCGCGGTTTCGAGAACATATGACGCGCTGTG	1860				
Qy	2393	ctgctcggcggtgacagcttcgggtggagcagacagttacagtttttgggggaaataaggaag	2452				







Db	1022	GTATGCTTGGATCTGGGATCCAAATTTTGAAGGCTGGATGAGATCTACCGAGTCGAGT	1081
Qy	1798	tacaagataaaattccgttgatgggttgatttagttccagttttccacagataaact	1857
Db	1082	TACAAGGATAAAATCCGTGGATGGTTCGATTTAGTGTTCAGATTTCCACAGATAACT	1141
Qy	1958	gcaggttgcatatattgtttaatgcatacgagattgaaccttgcggttcttaacagcta	1917
Db	1142	GCAGGTTTGCATATATATGTTAATGCCATCGAGATTGAACCTTGGCGCTTAAATCAGCTA	1201
Qy	1918	tatgctatgcaatatggtcacagtccctgtagttcatgaaactgggggacctccgagacaca	1977
Db	1202	TATGCTATGCAATAATGGTACAGTTCCTGTAGTTTCATGAACTTGGGGGCTCCGAGACACA	1261
Qy	1978	gtcgagaccttcaaccttttgggtcaaaagagagagggtcacagggtggggcttcttcca	2037
Db	1262	GTCCAGACCTTCAACCTCTTTGCTGTCAAAGCAGAGGAGGTACAGGTTGGCGCTTCTCA	1321
Qy	2038	ccgctaaccgtgacaaagattgttgggcaattgcgaaccggagatgcagaaattcaggagag	2097
Db	1322	CCGCTAACCGTGGACAAGATGTTGTGGGCATTTGCCAAACCGCGATGTCGACATTACGGGAG	1381
Qy	2098	cacaagcgttccttggagggggtcctatgaagcgaggcatgacgaagaacacatacgtggagac	2157
Db	1382	CACAAGCGTCTCTGGGAGGGCTCATATGAAGCGAGGCATGACGAAGACCATACGTGGGAC	1441
Qy	2158	catgccccgagcagtcacgacgagatcttcgagtgggcccttcgtggacaacacctacgtcca	2217
Db	1442	CATGCCCCGAGCAGTACGAGCAGATTTCTCGAGTGGGCGCTTCGTGGACCAACCTTACGTCA	1501
Qy	2218	tgtagacggggacctggggagggtccaaagtgcgagtcctctcagctctgaagacatcctct	2277
Db	1502	TGTAGACGGGACTGGGGAGGTTCCAAAGTCGAGTCTCTCTCAGCTCTGAAGACATCTCT	1561
Qy	2278	tcatctctcggggcccggaagataccocctgtacattgcgttgcctgctcacagttagag	2337
Db	1562	TCATCTCTTCGGCGGCGGAAGATACCCCTGTACATTTGCGTTGCTTGTCTACAGTAGAG	1621
Qy	2338	tcgcaatggcctgtgtgttggtttcggcttcgagaaacatatgacggctgtgctgct	2397
Db	1622	TCGCAATGCGCCCTGCTGTCTTTGGTTGCGCGGTTTCGAAACATATGACGGCTGTGCTGCT	1681
Qy	2398	gcggcgttgacagcttcgggttgagcacagttcacagttttggggaaataaggaaggagatgt	2457
Db	1682	CGCGCGTGACAGCTTTCGGGTGGACGACAGTTTACAGTTTGGGGAATAAGGAAGGATGT	1741
Qy	2458	gctgcaggatggttacaacagaagaccactcagatggcgcctctctgttcggtgttaca	2517
Db	1742	GCTGCAGATGGTTTAACAGCAAGACCACCTCAGATGCAGCCCTCTGTGTCGGTTTACA	1801
Qy	2518	gctgaactcagaaaccaactggtagctctttagccttagtgatgtgaagtctgttgccct	2577
Db	1802	GCTGAANTCAGAAACCACTGGTGACTCTTTAGCCTTAGTCATTGTGNAGTTTGTTCGCCCT	1861
Qy	2578	tctgtgtatgtgtgtcttgccttagctgacaaatatttgacctgttggaataatttatct	2637
Db	1862	TCTGTGTATGTTGCTCTGTCTTAGCTAGCTGACAAATATTTGACTCTTGGAGAAATTTATCT	1921
Qy	2638	tgtcgtctgtttttttaaatacaagagggggttccctcgattcttcatttaaaaaa	2697
Db	1922	TTGCTGCTGTTTTTTTTTAATCAAAAGAGGGGGTTTCTCTCGATTTTCAATTAATAAAAAA	1981
Qy	2698	aaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaa	2757
Db	1982	AA	2041
Qy	2758	aaaaaaaaaaaaa 2771	
Db	2042	AAAAAAAAAAAAA 2055	
RESULT	8		

A93359		A93359	2383 bp	DNA	PAT	22-JAN-2000
LOCUS		Sequence 1 from Patent WO9744472.				
DEFINITION		A93359				
ACCESSION		A93359.1	GI:6741623			
VERSION						
KEYWORDS		Zea mays.				
SOURCE		Zea mays.				
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.				
REFERENCE		Kossmann,J. and Froberg,C. NUCLEIC ACID MOLECULES CODING SOLUBLE MAIZE STARCH SYNTHASE Patent: WO 9744472-A 1 27-NOV-1997;				
AUTHORS		KOSSMANN JENS (DE); FROBERG CLAUS (DE)				
TITLE		Location/Qualifiers				
JOURNAL		1..2383				
FEATURES		/organism="Zea mays"				
source		/db_xref="taxon:4577"				
		/tissue_type="ENDOSPERM"				
		<2..1951				
CDS		/function="STARCH SYNTHESIS"				
		/codon_start=1				
		/product="SOLUBLE STARCH SYNTHASE"				
		/protein_id="CA869345.1"				
		/db_xref="GI:6741624"				
		/translation="ARGLLSLSAMTPSAVGAACLLARAPAAVAQDRAPRRLQRVLRRCVAELSRREGPAPRPMPALLAPLPVLPGFLPAEPTEGPALTPPVPDAGLGLGVGEPTAEGISINTVVVASQESEIIVGKEARAKVTQNI VEVTTGEASPYAKSGGLGDVCSGLPLAALRGHVMYPMRYLNAGTSDKANAYAFKTRIRPCFGEHEVTFFHEYEDVDWVFVDHPSHVRGNLXGDKEGAQNOPRYTLILCYANCEAPLVLELGGIYGQNMCVNDWHASLPVLAAKYREYGYKQRSLSLVTHNLHQGVEPASTYPDLGLCPNMYGALEMWFPEARRHALDKGEAVNFKGAADVADRIITVSKGSWEVTTABEGQLGNELLSSRSKSYLVNGVIDINDMNPATDKIPCHYSYDDLSGRKA KGAQLKGELGPIRDPLDTGIFGLRYQKDIIQII IIDLMDREDVOFMLSGDPLELDMMWRSTESIKDFKFWMGYFSVPVSHRTTAGCDIILLPSRFEPGCLNOLYAMOGTVPVVHATGLGRDIVTENFPGENGEQCTGWAFPLATTENMLWLTRLTAISTYREHKSWEGLMKRGMASKFTWDHAAEQYEIQFAWAFDIRPYVM"				
BASE COUNT		582 a 546 c 633 g 622 t				
ORIGIN						
		Query Match 50.3%; Score 1393; DB 6; Length 2383; Best Local Similarity 83.5%; Pred. No. 2.5e-178; Matches 1619; Conservative 0; Mismatches 310; Indels 10; Gaps 3;				
Qy		302 ccggggtcctgcgccccagcgctgcgcgcgatccggcgacaggccccggcgct 361				
Db						
Db		39 CCTCGCGGTGGGCCGCCTGCTCTCTCCGCGGGCGCGCTGGCCGCCGCTGCG				
Qy		362 ccgcctcgttcgc 421				
Db						
Qy		422 tcagcaggaggggcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 481				
Db						
Qy		482 tgccaggcttcctgc 541				
Db						
Qy		542 cgccccctgc 601				
Db						
Qy		602 ctgaggattccatgcagacgataatttgttggtcgaagtgcaggagttcttgagatcatgg 661				
Db						
Qy		662 atgcgaatgcagcaacctccaagctaaaagttaccagtagcatcgctgttgttgactcgttgaa 721				
Db						



Db 390 TTGGAAGAGGAGCAGCTCGAGCTAAAGTAACACAAAAACATTGCTTTGTAACTGGCGAAG 449  
QY 722 ctgctcctctatgcaaaagtcagggggttggaagatgtttgtgttgcgttaccaattgctc 781  
Db 450 CTTCCTCTTATGNAAGTCGCGGGTCTAGAGATGTTTGTGGTTTCATTCCCAAGTGTCTC 509  
QY 782 ttgctgctgtggttcaccgagtgatggttgtaataagcaagatacttaaaatgggtccctctg 841  
Db 510 TTGCTGCTGCTGTGTCACCGTGTGATGGTTGTAATGCCACAGATATTTAAATGGTACCTCCG 569  
QY 842 ataaaaactatgcaaaagcaattatacactcgaagcacataagattccatgctttggggg 901  
Db 570 ATAGAAATTATGCAATGCAATTTTACACAGAAAAACACATTCGATATTCATGCTTTGGCG 629  
QY 902 gatcacatgaagtgcacctttttcatgtagtagagacaacgctcagatgggtgtttgtcg 961  
Db 630 GTGAACATGAAGTTACCTCTTCCATGAGTATAGAGATTCAGTTGACTGGGTGTTTGTG 689  
QY 962 atcatcgtctatcatcagaccgggaagtttatatgagataattttgtgcttttgggtg 1021  
Db 690 ATCATCCCTCATATCAGACACCTGGAAATTTATATGGAGATAAGTTTGGTGTCTTTGGTG 749  
QY 1022 ataactcagttcagatacacactcttctgctatgctgcatcgagggcccaactaatccttg 1081  
Db 750 ATAACTCAGTTTCAGATACACACTCCTTTGCTTATGCTGCATGTGAGGCTCCTTTGGTCTTG 809  
QY 1082 aattgggagagatatattatggacagaattgcatgtttgttgaaagcattgcatgccca 1141  
Db 810 AATTGGGAGGATATATTATGGACAGATGTCATGTTTGTGTCATGATGGCANTGCCA 869  
QY 1142 gcttgtgcaagtcctctcttctgctgcaaaatatagaccatacgtgtttacagagattccc 1201  
Db 870 GTCTAGTGCCAGTCTTCTTGTGTCGAAAAATATAGACCATATGCTGTATTATAAAGACTCCC 929  
QY 1202 gcagcacccttgtatatacataatttagcacatcaggtgtgagcctgcaagtacatatc 1261  
Db 930 GCAGCATCTGTAAATACATAATTTAGCACATCAGGCTGTAGAGCTGCAGCACATATC 989  
QY 1262 ctgactggtggttgcctcctgaatgtagtgagcttttagaattgggtatttccagaatggg 1321  
Db 990 CTGACCTTGGTGTGCCACCTGAAATGGTATGAGCTCTGGAGTGGGTATTCCCTGAATGGG 1049  
QY 1322 caagggagcgtccctgtgacaaaggtgagcgagtttaactttttgaaaggagcagttgtga 1381  
Db 1050 CGAGGAGGATGCGCTGTGCAAGGGGTGAGCAGCTTAATTTTGAAGGTGCAAGTGTGA 1109  
QY 1382 cagcagatcggattgtgacctgacgtcaggtgttattcatggtgaggttcacaaactgctgaag 1441  
Db 1110 CAGCAGATCGAATCGTGACTGTGAGTAAAGGTTATTCATGGGAGGTCAACAATGCTGAAG 1169  
QY 1442 gtgacagggcctcaatagactcttaagctcccgaaaaagtgtattgaaatggattgtaa 1501  
Db 1170 GTGGACAGGGCCCTCAATGAGCTTAAAGCTCCAGAAAGAGTGTTAAACGGGAATGTAA 1229  
QY 1502 atggaattgaattatgattgaaaccccccacacagacaagtgctcctccatcatatt 1561  
Db 1230 ATGGAATTGACATTAATGATGGAACCCCTGCCACACACAATGATGCCCTGTGCATATT 1289  
QY 1562 ctgtcgtgacacctctctggaagggccaaatgtaaagctgaattgacagaggaggtgggtt 1621  
Db 1290 CTGTTGATGACCTCTCTGGAAGGCCCAATGTAAAGTGTCATTCAGAGAGGAGCTGGGT 1349  
QY 1622 taccgtgaagggaggatgttctcctgattggtcttatttgaagactggattaccagaaag 1681  
Db 1350 TACCTATAAGCCCTCTGTTCTCTGATTTGCTTATTTGGAAGATTGGATTATCAGAAAG 1409  
QY 1682 gcaatcatctataaaatggccattccagagctcatgagggagcgctgcgaatttgta 1741  
Db 1410 GCATTGATCTCATCAACCTTATCATACCAAGATCTCATGCGGGGAAGATGTTCAATTTGTCA 1469  
QY 1742 tgtttggtatctgggtcccaatttttgaagggctggatgagatctaccaggtcaggttaca 1801  
Db 1470 TCGTGTGATCTGCTGACCCAGAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCA 1529

QY 1802 aggataaaattccatggatgggttggtatttagtttagttccagtttccccacagaataactgcag 1861  
Db 1530 AGGATAAAATTTCTGGATGGTGGATTTAGTTTCCAGTTTCCACCCGAATAACTGCCG 1589  
QY 1862 gtgcgatatattgttaatgccatccagatttgaacctctgcgtcttaatacagctatatg 1921  
Db 1590 GCTGCGATATATTGTAATGCCATCCAGATTCCGAACCTTGTGCTCAATCAGCTATATG 1649  
QY 1922 ctatgcaaatatgttacagtttccctgttagttcatggaactgggggcccctccagagacacagtcg 1981  
Db 1650 CTATGAGTATGTCACAGTTCCTGTTGCTCCATCAACTGGGGGCTTAGAGATACCGTGG 1709  
QY 1982 agaccttcaaccttttgggtgcaaaaggagagagaggttacaggttgggctttctcaccgc 2041  
Db 1710 AGAATTTCAACCCCTTTCCGGTGAGATGGAGAGCAGGGTACAGGGTGGGCATTGCGACCCC 1769  
QY 2042 taacctgtgacaagaatgtgtggcattgcgaaccgcgattgcgacattcagggagacaca 2101  
Db 1770 TAACCACAGAAAAACATGTTGTGCACATTGGGAATGCAATATCTACATACAGGGAACACA 1829  
QY 2102 agcgtctctgggggggctcatgaagcaggcatgacgaagaagaccatacctggggaccatg 2161  
Db 1830 AGTCCTCTCTGGGAAGGCTTAATGAAGCAGGATGTCAAAAGACTTTCACGTGGGACCATG 1889  
QY 2162 cc-ccagcagctacgacagatcttcgagtggtggccttcgtggaccaacctacatcgt 2220  
Db 1890 CCCTGAACANTACGAACAAATCTTCCAGTGGGCTTTCATCGATCGACCCCTATGTCTATGT 1949  
QY 2221 agacgggggactggggaggt 2239  
Db 1950 AAAAAAGGCCAACAAAGTGGT 1968

RESULT 9  
AF036891  
LOCUS  
DEFINITION  
Zea mays starch synthase I precursor (Ssl) mRNA, nuclear gene  
ACCESSION  
AF036891  
VERSION  
AF036891.1  
KEYWORDS  
GT:2828011  
SOURCE  
Zea mays.  
ORGANISM  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
1 (bases 1 to 2991)  
AUTHORS  
Knight,M.E., Harn,C., Lilley,C.E.R., Guan,H., Singletary,G.W.,  
Mu-Forster,C., Wasserman,B.P. and Keeling,P.L.  
TITLE  
Molecular cloning of starch synthase I from maize (W64) endosperm  
and expression in Escherichia coli  
JOURNAL  
Plant J. 14 (5), 613-622 (1998)  
MEDLINE  
98340555  
PUBMED  
9675904  
REFERENCE  
2 (bases 1 to 2991)  
AUTHORS  
Knight,M.E., Harn,C., Lilley,C.E.R., Guan,H.P., Singletary,G.W.,  
Mu-Forster,C., Wasserman,B.P. and Keeling,P.L.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (03-DEC-1997) ExSeed Genetics, 1568 Food Science  
Building, ISU, Ames, IA 50011, USA  
FEATURES  
Location/Qualifiers  
1..2991  
/organism="Zea mays"  
/strain="W64A"  
/db\_xref="taxon:4577"  
/map="9; between npi266 and wx1"  
/chromosome="9"  
/tissue.type="endosperm"  
1..2991  
/gene="Ssl"  
602..2470  
/gene="Ssl"















[illegible]



Db 354 CCGACCCCGCGCGTTCGCGGATCTCGGCGTGGGGAGATCGAGCCGATCTAAGAGGT 413  
Qy 598 attgctgagattccatcgacagacataattgtggtcgtcaagctgagcaggattctgagatc 657  
Db 414 CTCACAGAAGATCCATCGCAACAAACAAATTTTGTGGCTAGTGAGCAGGAGCTGAGATC 473  
Qy 658 atggatcggaatgagcaaacctcaagctaaagtctacagctagcagctggtttgtgactgggt 717  
Db 474 ATGGATGTGAAGGAGCAAGCTCAAGCTAAAGTAAACGACGACGCTGTGTCTGTAAACCGGT 533  
Qy 718 gaagctctcttatgcaagtcagggttggaagatgtttgtggttctggttaccatc 777  
Db 534 GAAGCTTCTCTTATGCAAAATGACAGTGGACTAGGAGATGTTGTGGTTCTCACTGCCAAAT 593  
Qy 778 gctcttctgctggtgcacccgagtgatggttgaatgccaagatacttaaatgggtcc 837  
Db 594 GCTCTTCTCTCGTGTGTCATCGTGTGATGTTGTAATGCCGAGATACATGAACGGGGCC 653  
Qy 838 tctgataaaactatgcaagggcaatttcaactgcgaagcacataagattccatgcttt 897  
Db 654 TTGAACAAAATTTTGGAAACGCACTTTTACACTGAGAGACACATTAAGATTCCATGCTTT 713  
Qy 898 ggggatcacatgaagtcagctttttcatgagctatagacacacgcgattgggttt 957  
Db 714 GCGGAGAACATGAAGTTACTTTTTTTCACAGATATAGGGATCTGTGTGATGGGTGTTT 773  
Qy 958 gtcgataatccgtatcatcacagaccaggaatttlatgagagataattttggtgctttt 1017  
Db 774 GTTCATCATCCCTCATATCATAGACCTGGAAATTTGTATGGAGATAATTTTGGTGTCTT 833  
Qy 1018 ggtgataatcaagtcagatacaactccttctgctatgctgcagcgaggcccaactaatc 1077  
Db 834 GCGGATAATCAGTTTCAGATACACACTCTGTCTATGCGCGTGTGAAGCCCGCATTAAT 893  
Qy 1078 ctgaaattggagagatatttattgacagaaattgcatgtttgttgaacgattgcat 1137  
Db 894 CTGGAACCTGGGAGGATATCTATGGACAGAAATGCATGTTTGTGTGAATGATGGCAT 953  
Qy 1138 gccagccttggcagctccttctgctgcaaaatatagaccatacagctgttttacagagat 1197  
Db 954 GCCAGTCTGTGCCAGTCCCTTCTGTCTGCAAAATATAGACCATATGTTGTTCACAGGAT 1013  
Qy 1198 tcccgagcacccttggtatacataatttagcacatcagggtgtggagcctgcaagtcata 1257  
Db 1014 GCGCGCAGTCTTGTTCATACATAATCTAGCACATCAGGCTGTGGAGCCTGCCAGTACA 1073  
Qy 1258 tatectgactggatgctcctggaatgtatgagctttagaatgggtatttccagaa 1317  
Db 1074 TATCTGACCTGGGATGGCCACCTGAATGATGAGACATAGAAATGGGTGTTTCCAGAG 1133  
Qy 1318 tgggcaaggaggatgcctcttgacaaagggtgagcagcttaacttttgaagggagcagtt 1377  
Db 1134 TGGCAAGGCGCATGCTCTTGCAAGGCTGAGCAGCTCAATTTTAAAGGCGCAGTT 1193  
Qy 1378 gtgacagcagatcggaattgtgacctgacgctcagtgaggttattcattgaggggtcacaactgct 1437  
Db 1194 GTGACAGCAGATCGAATGTGACTGTGACGACGAGGGTATTTCATGGGAGGTGACAACTGCT 1253  
Qy 1438 gaaggtgacagggcctcaatgactcttaagctcccgaaagaaagtgtatgaatggaatt 1497  
Db 1254 GAAGGTGGGAAGGCGCTCAATGAGCTCTTAAGCTCCCGAAGAGTGTATTGAATGGAAAT 1313  
Qy 1498 gtaaatggaattgacataatgattggaaccccccacacagacagtgctctccctcatcat 1557  
Db 1314 GTAATGGAATTGACATTAATGATTGGACCCATCCACAGACAGTTTCTCCCTTATCAT 1373  
Qy 1558 tattctgcatgacacctctctggaagggccaaatgtgaaagctgaattgcagaaggagttg 1617  
Db 1374 TATTCTGTTGATGACCTGTCCGGAAGGCCAAAGTGTAAAGCTGAATTCAGAGAGGAGCTG 1433  
Qy 1618 ggttacctgaaggaggaggttctctctgattggttttattggaagcagattaccag 1677

Db 1434 GGTTTACTATAAGCCCGATGTGCCCTCTGATTCGGCTTTATTGGGAAGATTGGACTATCA 1493  
Qy 1678 aaagcattgatctcattaaatgcccattccagagctcagtgaggagcagtgcaattt 1737  
Db 1494 AAAGCATTTGATCTAATTAACCTTGGCATTCAGATCTCATGCGGACAAATATTCAATTC 1553  
Qy 1738 gtcgtgttgatctggggatcccaatttttgaaggtcggatgagatctaccaggtcaggt 1797  
Db 1554 GTCATGTTGGATCTGGTGACCCAGGTTTGAAGGATGGATGAGATCCACAGAATCAGGG 1613  
Qy 1798 tacaagataaaatccgtggaatgggttggatttagttccagtttcccacagataaact 1857  
Db 1614 TACAGGATAAATTTCTGTGGATGGGTGGATTTAGTGTTCAGTTTCCCACCAATRACT 1673  
Qy 1858 gcaggttgcgataattgttaattcccatcgagatttgaaccttcggtcttaatacagcta 1917  
Db 1674 GCAGTTTGGGATATATTGTTGATGCCATCCAGATTCCGAACCTTGTGGCTCAATCAGCTA 1733  
Qy 1918 tatgctatgcaatgtgtacagtttccctgtagtttccatggaacctggggcctccgagaaca 1977  
Db 1734 TATGCTATGCAATATGTTACAGTGCCTGTTGTTTCATGGAATGGAGGCTCCAGAGATA 1793  
Qy 1978 gtcgagaccttcaaccttttgggtcaaaagagagaggtacaggggtggcgcttctca 2037  
Db 1794 GTGAGAAATTTTAAACCGTTCCTGTGAGAAAGAGAGGAGGTACAGGCTGGGATTCG 1853  
Qy 2038 ccgctaacc-gtggacaagattgttggcatttgcgaacccgcgagatgcgacattcaggga 2096  
Db 1854 CCATAACCATTTGAAAAAATGCTGTGGCATTCGGGATTCGGAATTCGACATACAGGA 1913  
Qy 2097 gcacaaagcctcctggagggggtcattgaaagcagggatgacgaaagaccatacagtgga 2156  
Db 1914 ACACAAGTCTCTTGGGAGGCTCTAATGAAGCAGGAGTGTCAAGCGACTTTTACATGGGA 1973  
Qy 2157 ccatgccc-ccgagcagctacgacagatcttctgagtcggttcgtgacacacccctcagt 2215  
Db 1974 CCATGCGGCTCACAGTATGACAGATCTTCGAATGGGCTTCATGATCAACCATATGT 2033  
Qy 2216 catgtagacgg 2226  
Db 2034 CATGTAATGG 2044

## RESULT 13

RICSSS Rice mRNA for soluble starch synthase, complete cds. 01-FEB-2000  
LOCUS D16202.1 GI:450484  
DEFINITION soluble starch synthase.  
ACCESSION D16202.1  
VERSION 1  
KEYWORDS soluble starch synthase.  
SOURCE Oryza sativa  
ORGANISM Oryza sativa  
REFERENCE 1 (bases 1 to 2533)  
AUTHORS Baba,T., Nishihara,M., Mizuno,K., Kawasaki,T., Shimada,H., Kobayashi,E., Ohnishi,S., Tanaka,K. and Arai,Y.  
TITLE Identification, cDNA cloning, and gene expression of soluble starch synthase in rice (Oryza sativa L.) immature seeds  
JOURNAL Plant Physiol. 103 (2), 565-573 (1993)  
MEDLINE 94302151  
REFERENCE 2 (bases 1 to 2533)  
AUTHORS Baba,T.  
TITLE Direct Submission  
JOURNAL Submitted (06-MAY-1993) to the DDBJ/EMBL/GenBank databases. Tadashi Baba, University of Tsukuba, Institute of Applied Biochemistry, Tennohdai 1-1-1, Tsukuba Science City, Ibaraki 305, Japan (Tel:298-53-6632, Fax:298-53-6632)  
COMMENT Submitted (06-May-1993) to DDBJ by: Tadashi Baba  
Institute of Applied Biochemistry  
University of Tsukuba











[illegible]



Qy 906 acatgaagtgaaccttttttccatgaagtatagagacaacgtcgagattgggtgttggatcatca 965  
Db 1208 ACATGAAGTTACCTTCTCCATAGTATAGAGATTCAGTTCACGGGTGTTGTGTATCA 1267  
Qy 966 tccgtcatatcacagacagaaagtcttatatggagataaatttttggctcttttggatgataa 1025  
Db 1268 TCCCTCATATCACAGACCTGGAATTTATATGGAGATTAAGTTTGGTGTCTTTGGTGATAA 1327  
Qy 1026 tcagttcagatcacacctcttctgtatgtctgcatcgagagcccccactaatacctttgaatt 1085  
Db 1328 TCAGTTTCAGATACACACTCCTTCTCTATGCTGCATGTGAGGCTCCTTTGGTTCCTTGAATT 1387  
Qy 1086 gggaggatataatttatggacagaattcagatgttggttgtgaacagattggatgcagacct 1145  
Db 1388 GGGAGGATATATTTATGGACAGAATTGCATGTTGGTGTGTCGAATGATTGGCATGCCAGTCT 1447  
Qy 1146 tgtgccagtccttctctcacaataatagaccatcaggtgttttacagagattcccgag 1205  
Db 1448 AGGCCAGTCTCTTCTGCTCAAAATATAGCCATATGGTGTATTAAAGACTCCCGCAG 1507  
Qy 1206 cacccttgttatacataatttagcacatcaggggtgtgagcctgcaagtacatatcctga 1265  
Db 1508 CATTCCTGTAATACATAAATTAGCACATCAGGGGTGAGAGCCTGCAAGCACATATCCTGA 1567  
Qy 1266 tctgggattgcctcctcctgaattggtatgagatttagaattgggtatttccagaatgggcaag 1325  
Db 1568 CCTTGGGTTGCCACCTGAATGGTATGGAGCTGAGAGTGGGTATTCCTGAAATGGGCGAG 1627  
Qy 1326 gaggcatgcccttgacaaggtgagggcagtttaactttttgaaaggagcagttgtgacagc 1385  
Db 1628 GAGGCATGCCCTTGACAAGGGTGAGGCGAGTTAATTTTTGAAAGTGCAGTTGTGACAGC 1687  
Qy 1386 agatcggattgtgacctcaggtcaggtgtattcatggagggtcacaaactcgtgaaggtg 1445  
Db 1688 AGATCGAATCGTGACTGTCACTAAGGGTTATTCTGGAGGTGCACAACTGCTGAAGGTGG 1747  
Qy 1446 acagggcctcaatgagctcttaagctcccgcaaaaaagtgtattgaattgtaaatgg 1505  
Db 1748 ACAGGGCTCAATGAGCTCTTAAGCTCCGAAAGAGGTATTAACCGGAATTTGTAATGG 1807  
Qy 1506 aattgacattgaattggaacccacacagacagaagtgtctcctcctcatcattctgt 1565  
Db 1808 AATTGACATTAAATGTTGGAACTCCGACACAGACAATGTATCCCTGTCTATTATTCTGT 1867  
Qy 1566 cga tgaacctctctggaagccaaatgaaagctgaaatgcagaagaggttgggtttacc 1625  
Db 1868 TGATGACCTCTCTTGAAGGCTAAATGTAAAGGTGCATTGCAGAAGGAGCTGGGTTTACC 1927  
Qy 1626 tgaagggagagatttctcctcctgattggctttatttggagagactggattaccagaaggcat 1685  
Db 1928 TATAAGGCTGTATGTTCTGTATTGGCTTTATTGGAGATTTGGATTATCAGAAAGGCAT 1987  
Qy 1686 tgatctcatlaaaatggccattccagagctcagatgagggagacgtgcaatttgcattgct 1745  
Db 1988 TGATCTCATTCACCTTATCATACCAGATCTCATCGGAAGAAATGTTCAATTTGTCTATGCT 2047  
Qy 1746 tggatctggggatcccaatttttgaaggctggatgagatctaccagtcgaggttacaagga 1805  
Db 2048 TGGATCTGGTGACCCAGAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCAAGGA 2107  
Qy 1806 taaattccgtgagattggattgattggttccaggtttccacagagaataaactgcaggttg 1865  
Db 2108 TAAATTTGGTGGATGGGTTGGATTTAGTGTTCAGATTTCGCCCAAGATAACTGGCGCTGG 2167  
Qy 1866 cgatatattgttaatgccaatgagatttgaaccttgcggtctttaaactcagctatatgctat 1925  
Db 2168 CGATATATTGTTATGCCATCCAGATTCGRACCTTGTGGTCTCAATCAGCTATATGCTAT 2227  
Qy 1926 gcaataatggtacagttcctcgtgagttcattgaaactgggggctcccgagacacagtcgagac 1985  
Db 2228 GCAGTATGGCACAGTTCCTGTTGTCCTGTCGAACCTGGGGGCTTAGAGATACCGTGGAGAA 2287  
Qy 1986 cttcaaccttttgggtgcaaaaggagagggttacagggtggggttctcaccgcgctaac 2045

Db 2288 CTTCAACCCCTTTCGGTGAGAATGGAGAGCAGGTTACAGGGTGGGCATTCGACCCCTAAC 2347  
Qy 2046 cgtggacaagatg-ttgtgggcatctgcgaaccgcgcatgtcgacatttcaggaggagcacaagc 2104  
Db 2348 CACAGAAAACATGTTTGTGGACATTCGAACTGCAATATCTACATACAGGGAACACAAGT 2407  
Qy 2105 cgtcctggaggggcctcatgaagcagagcatgacgaaagacca taccgtgggaccatgcc- 2163  
Db 2408 AATAATGGGAAGGCTAATGAAGCCAGGCATGTCAAAAGAGTTTCACGTGGGACCATGCCG 2467  
Qy 2164 ccgagcagttacgagcagcatcttccaggtgggcttctcgtgg 2202  
Db 2468 CTGAACAATACGAACAAATCTTCAGTGGGCTTCATCG 2506

Search completed: March 28, 2002, 14:54:59  
Job time: 9898 sec







**THIS PAGE BLANK (USPTO)**

---







**THIS PAGE BLANK (USPTO)**

---



QY 1313 cagaatgggcaaggagcagctgacaaagggtgagcagcttaacttttgaaggag 1372  
DB 781 cagaatgggcaaggagcagctgacaaagggtgagcagcttaacttttgaaggag 840  
QY 1373 cagttgtgacagcagatcgagattgtgacgctcagtcaggttattcatggtgaggtcacaa 1432  
DB 841 cagttgtgacagcagatcgagattgtgacgctcagtcaggttattcatggtgaggtcacaa 900  
QY 1433 ctgctgaagggtgacagggccctcaatgagctcttaagcttcccgaaagtgtattgaatg 1492  
DB 901 ctgctgaagggtgacagggccctcaatgagctcttaagcttcccgaaagtgtattgaatg 960  
QY 1493 gaattgtaaatggaattgacattaaatggaacccaccacagacagagtgctccctc 1552  
DB 961 gaattgtaaatggaattgacattaaatggaacccaccacagacagagtgctccctc 1020  
QY 1553 atcattattctgtcgatgacctctctgaaagcccaaatgtaaaagctgaattgcagaagg 1612  
DB 1021 atcattattctgtcgatgacctctctgaaagcccaaatgtaaaagctgaattgcagaagg 1080  
QY 1613 agttgggtttacgttaaggagagtgcttctctgattggtcttattggaagactggatt 1672  
DB 1081 agttgggtttacgttaaggagagtgcttctctgattggtcttattggaagactggatt 1140  
QY 1673 accagaaaggcattgatctcattaaatgcccattccagagctcatgagggagacgtgc 1732  
DB 1141 accagaaaggcattgatctcattaaatgcccattccagagctcatgagggagacgtgc 1200  
QY 1733 aattgtcatgctgtgattggtggtcccaatttttgaaggctggtgagatctaccaggt 1792  
DB 1201 aattgtcatgctgtgattggtggtcccaatttttgaaggctggtgagatctaccaggt 1260  
QY 1793 cgagttacaaggataaattccgtggtggtggtggtggtggtggtggtggtggtggtggt 1852  
DB 1261 cgagttacaaggataaattccgtggtggtggtggtggtggtggtggtggtggtggtggt 1320  
QY 1853 taactgcaggttcgcatatattgtaacccatcgagattgaaaccttgcggtcttaact 1912  
DB 1321 taactgcaggttcgcatatattgtaacccatcgagattgaaaccttgcggtcttaact 1380  
QY 1913 agctatatgctatgcaatggttacagttcctgtagttcatggaactggggcctccgag 1972  
DB 1381 agctatatgctatgcaatggttacagttcctgtagttcatggaactggggcctccgag 1440  
QY 1973 acacagtcagaccttcaaccttttggcgcaaaaggagaggtgacaggtggggtggtggt 2032  
DB 1441 acacagtcagaccttcaaccttttggcgcaaaaggagaggtgacaggtggggtggtggt 1500  
QY 2033 tctcacgcctaacctgggcaagatgtgtggcattgcgaacccgagatgcgacattca 2092  
DB 1501 tctcacgcctaacctgggcaagatgtgtggcattgcgaacccgagatgcgacattca 1560  
QY 2093 gggagcacagcgtcctggtggaggggtcattgaagcgagggcatgacgaaagacattacgt 2152  
DB 1561 gggagcacagcgtcctggtggaggggtcattgaagcgagggcatgacgaaagacattacgt 1620  
QY 2153 gggagcacatccccagcagcagcagcagatttctgagtggtggccttgcgtgacaccccta 2212  
DB 1621 gggagcacatccccagcagcagcagcagatttctgagtggtggccttgcgtgacaccccta 1680  
QY 2213 cgtcatgtagaggggactgggaggttccaagtgcagtgctccttcagctctgaaagacat 2272  
DB 1681 cgtcatgtagaggggactgggaggttccaagtgcagtgctccttcagctctgaaagacat 1740  
QY 2273 cctcttcaccttcccgccgagggaggtacccctgtacattgctgttctcgtctacag 2332  
DB 1741 cctcttcaccttcccgccgagggaggtacccctgtacattgctgttctcgtctacag 1800  
QY 2333 tagagtcgaaatgcgctgcttctgttctgttccggttccgagacatagcgggtgtg 2392  
DB 1801 tagagtcgaaatgcgctgcttctgttctgttccggttccgagacatagcgggtgtg 1860  
QY 2393 ctgctggggcgtgacagcttccggttggagcagagttacagtttggggaataaggagag 2452

DB 1861 cgtgctgcgggtgacagcttccggtggaacagttacagtttggggaataaggagag 1920  
QY 2453 gatgtgctcagagtggttaacagcaaaagcaccactcagatggcagctctctgtccgtg 2512  
DB 1921 gatgtgctcagagtggttaacagcaaaagcaccactcagatggcagctctctgtccgtg 1980  
QY 2513 ttacagctgaatcagaacccaactgggtgactctttagcttagtgattgtagtgaagtgtt 2572  
DB 1981 ttacagctgaatcagaacccaactgggtgactctttagcttagtgattgtagtgaagtgtt 2040  
QY 2573 tgctctctgtatgttctgtcttcttagctgacaaaattttgacctgttgagaaattt 2632  
DB 2041 tgctctctgtatgttctgtcttcttagctgacaaaattttgacctgttgagaaattt 2100  
QY 2633 tatcttgcgtgctgtttttttaaatacaaaagaggggttctcctcgatttcatataaaa 2692  
DB 2101 tatcttgcgtgctgtttttttaaatacaaaagaggggttctcctcgatttcatataaaa 2160  
QY 2693 aaaaaaatacaaaagaggggttctcctcgatttcatataaaa 2752  
DB 2161 aaaaaaatacaaaagaggggttctcctcgatttcatataaaa 2220  
QY 2753 aaaaaaatacaaaagaggggttctcctcgatttcatataaaa 2771  
DB 2221 aaaaaaatacaaaagaggggttctcctcgatttcatataaaa 2239

## RESULT 4

AAT95785  
ID AAT95785 standard; cDNA to mRNA; 2383 BP.

XX AC AAT95785;

XX DT 22-MAY-1998 (first entry)

XX DE Maize starch synthase type I cDNA.

XX KW Maize; starch synthase type I; starch; ds.

XX OS Zea mays.

XX FH Key Location/Qualifiers

FT CDS 2..1951

FT /\*tag= a

XX /\*product= starch\_synthase\_type\_I

PN DE19619918-A1.

XX PD 20-NOV-1997.

XX PF 17-MAY-1996; 96DE-1019918.

XX PR 17-MAY-1996; 96DE-1019918.

XX PA (PLAN-) PLANTTEC BIOLOGIE GMBH.

XX PI Froberg C. Kossmann J;

XX DR WPI: 1998-000821/01.

XX DR P-PSDB; AAW38218.

XX PT DNA encoding maize starch synthase type I protein - for producing transgenic plants

XX PS Claim 1; Pages 16-20; 23pp; German.

XX CC The present sequence encodes maize starch synthase type I, useful in the production of starch. Starch can be used in various conventional starch applications, e.g. starch hydrolysate products, foods, papermaking, adhesives, textiles, building materials, soil stabilisation, agrochemicals, pharmaceuticals, cosmetics, coal briquettes, ore and coal slurries, foundry casting, rubber, leather



**THIS PAGE BLANK (USPTO)**

---



```

Db 1861 CTGCTGCGCGGTGACACCTTCGGGTGGACGACAGTTACAGTCTTTGGGAATAAGGAAGG 1920
Qy 2453 gatgtgctgaggtggttaacagcaagcaccactcagatggcagcctctctgctcgtg 2512
Db 1921 GATGTGTCAGGATGTTTAAACAGAAAGCACCACACTCAGATGGCAGCCTCTCTGTCGCTG 1980
Qy 2513 ttacagctgaatcagaaacaaactggctgactctttgacctttagtctgagttgtg 2572
Db 1981 TTACAGCTGAATCAGAAACAACTGGTGACTCTTTAGCCTTAGTATTGTAAGTTGT 2040
Qy 2573 tgccttctgtgtgtgtgtctctctcttagctgacaaatattgacctgttgagaaattt 2632
Db 2041 TGCTCTCTGTGTATGTGTCTCTCTAGCTGACAAATATTTGACCTGTGAGAAATTT 2100
Qy 2633 tatcttctgtctgttttttttaatacaaaagggggttctctcagatttcattaaaa 2692
Db 2101 TATCTTTGCTGCTGTTTTTTTTTAATCAAAAGAGGGGGTTCCTCCGATTTCAATTAATA 2160
Qy 2693 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2752
Db 2161 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2220
Qy 2753 aaaaaaaaaaaaaaaaaaaaaa 2771
Db 2221 AAAAAAAAAAAAAAAAAAAAAA 2239

RESULT 7
TAU48227
LOCUS Triticum aestivum 2055 bp mRNA PLN 12-JUN-1996
DEFINITION Triticum aestivum soluble starch synthase mRNA, partial cds.
ACCESSION U48227.1 GI:1373149
VERSION U48227.1
KEYWORDS
SOURCE wheat.
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 2055)
Block, M., Loerz, H. and Luetticke, S.
Direct Submission
Submitted (01-FEB-1996) Martina Block, University of Hamburg,
Institute of General Botany, Centre of Applied Molecular Biology,
AMP II, Ohnhorststr. 18, Hamburg, 22609, Germany
On Jun 12, 1996 this sequence version replaced gi:1335887.
Location/Qualifiers
1..2055
/organism="Triticum aestivum"
/cultivar="T.a., Florida"
/db_xref="taxon:4565"
/tissue_type="endosperm of kernels"
<1..1834
/codon_start=2
/product="soluble starch synthase"
/protein_id="AAB02197.1"
/db_xref="GI:1373150"
/translation="EAPYAKSGGLGVDCYGLPALTAAARGHVRVWVMPRYLNGSSDKN
YAKALYTAHKIKPCFGGSHVTFHEYRDVDFVDPHSYHRPGLSDYGNFGAFGD
NOFRYTLICYAACEAPLILEGGYIYGNCFMVNDWHASLVPLVLAAYRPPYGVYRD
AVVTADRTVTSQYQSWETVTAEGGGLNELLSRKSVLNGIVNGIDINDNPTDKC
LPHYSVDDLSGAKKAEQLKELGPLVEDVPLIGFGLRDYQKGLDLKWLAPELM
EPDGNOLYAMQYGVTPVWHTGSLRDTVEFPNFGAKGEGGTWAP SPLTVDKMLWA
LRTAMSTFRHKPSMEGLMKRGMKDHWDHAPSSTSRSSGSPSWNTPTSCRRGLGRS
KCESPSALKYSSSFRPEGPCTLRCPATVSEQCACLLWFAAGSTYDGCACAAAVTAS
GGROLPWGIKKCAAGWLAKHSDGSLSVRVTAERINLVL"
BASE COUNT 570 a 399 c 534 g 552 t
ORIGIN

```

Query Match 74.1%; Score 2054; DB 8; Length 2055;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-267;  
 Matches 2054; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 718 gaagctgctccttatgcacaaagtcagggggttgaggagatgtttgttgcgttaccacatt 777
Db 2 GAAGCTGCTCTTATGCAAAAGTCAGGGGGTTCGGGAGATGTTTGTGGTTCGTACCAATT 61
Qy 778 gctcttgcctcgttggtcacocagtgatggttgtaatgccaaagatacttaaatgggtcc 837
Db 62 GCTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
Qy 838 tctgataaaactatgcacaaagtcacattatcacctgcgaagcacaattgaattccatcttt 897
Db 122 TCTGATAAAACTATGCAAAAGCATTTATACACTGCGAAGCACATTAAGATTCCATCTTT 181
Qy 898 ggggagatcacatgaagtgaccttttttcatgagatagagacaaagtcagtggtggttt 957
Db 182 GGGGGATCACATGAAGTGACCTTTTTCATGAGTATAGAGACAACTGCGATTGGGTGTTT 241
Qy 958 gtcgataccgctcattatcacagaccaggaagtattataggagataatttgggtgctttt 1017
Db 242 GTCGATCATCCGCTCATATCACAGACCAGGAAGCTTTATATGAGAGATAATTTTGGTCTTT 301
Qy 1018 ggtgataatcagttcagatacacactcctttgctatgctgcagcgagcccccactatc 1077
Db 302 GGTGATAATCAGTTCAGATACACACTCTCTTGTATGCTGCTGATGCGAGGCCACCTAATC 361
Qy 1078 cttgaattgggagatattatttgacagaaatgacatgctgttggtggaagattggcat 1137
Db 362 CTTCAATTGGGAGGATATATTATGGACAGAAATTTGCATTTGTTTGTGTAACGATTGGCAT 421
Qy 1138 gccagccttgcagcctccttctgctgcaaaatagaccatacgcgtgtttacagagat 1197
Db 422 GCCAGCCTTGGCAGCTCTTCTGCTGCAAAATATAGACCATACGGTGTTCAGAGAT 481
Qy 1198 tccgcagcacccttcttatacataatttgacacatcagggtgtggagcctgcgaagta 1257
Db 482 TCCGCGACGACCTTGTATACATAAATTTAGCACATCAGGGTGTGGAGCTGCAAGTACA 541
Qy 1258 tctcgtatgctgggtgctcctcctgaatggtgagctgttagaagggttattccagaa 1317
Db 542 TATCCTGATCTGGGATTCCTCTCTGATGCTATGGAGCTTTGAATGGGTATTTCCAGAA 601
Qy 1318 tggcgaagagcagatcccttgcacagggtgagcgagtttaacttttgaaggagcagtt 1377
Db 602 TGGCAGAGGAGGATCCCTTGCACAGGGTGGAGCATTAACCTTTTGAAGAGGAGCAGTT 661
Qy 1378 gtgacagcagatcggtatgtgacccgtcagtgaggttatctatgaggtgcacaactgct 1437
Db 662 GTGACAGCAGATCGGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721
Qy 1438 gaagtgacagggcctcctcaatgagctcttaagctcccccagaaagtgtctcctcatcat 1497
Db 722 GAAGTGGACAGGGCCTCAATGAGCTCTTAAGCTCCCGAAAAGGTATTTGAATGGAAAT 781
Qy 1498 gtaaatgaattgacatttaagtggaaagccacacagacaaagtgtctcctcatcat 1557
Db 782 GTAATGAATTTGACATTAATGATTTGAACCCACACACAGACAACTGCTCCTCTCATCAT 841
Qy 1558 tttctgtcgtatgacctctctggaagcccaaaagtgaagtgaattgcagaaggttg 1617
Db 842 TATTCTGCGATGACCTCTCTGGAAGGCCAAATGTAAGCTGAATTCAGAAAGGAGTTG 901
Qy 1618 gtttacctgtaaggaggtatgttctcctgattggtcttattggaagactggattaccag 1677
Db 902 GGTTTACTGTGAAGGAGGATGTTCTCTCTGATTGGCTTTATTTGGAAGACTGGATTACCAG 961
Qy 1678 aaaggcattgactctcattaaatggccattccagagctcatgagggagagcgtcaattt 1737
Db 962 AAGGCATTGATCTCATTAATAATGGCCATTCACAGAGCTCATGAGGGAGGACGTCAATTT 1021
Qy 1738 gtcagctgtggtctctgggttcccaatttttgaaggctggtgatgatctaccagtgagt 1797

```



**THIS PAGE BLANK (USPTO)**

---



```

A93359
LOCUS       A93359             2383 bp     DNA
DEFINITION Sequence 1 from Patent WO9744472.
ACCESSION  A93359
VERSION    A93359.1  GI:6741623
KEYWORDS
SOURCE
ORGANISM   Zea mays.
            Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 2383)
            Kossmann, J. and Froberg, C.
NUCLEIC ACID MOLECULES CODING SOLUBLE MAIZE STARCH SYNTHASES
PATENT: WO 9744472-A 1 27-NOV-1997;
KOSMANN JENS (DE); FROBERG CLAUS (DE)
FEATURES
            Location/Qualifiers
                1..2383
                /organism="Zea mays"
                /db_xref="taxon:4577"
                /tissue_type="ENDOSPERM"
                <2..1951
                /function="STARCH SYNTHESIS"
                /codon_start=1
                /product="SOLUBLE STARCH SYNTHASE"
                /protein_id="CAB69545.1"
                /db_xref="GI:6741624"
                /translation="ARGLSLSAMTPSAVGAACLLLAARAANPAAVGDRARPRLQRY
                LPRCAVLSRSGPAPRPPAPALLAPLPVGFLEAPAEPTGEPALTPPPVDPAGLGLV
                GVPEGIAGSDINTVVVAASEDSEIWEQEARAKVTONTIVFTGEASPIAKSGGLG
                DVGSLPVALAARHVMVMPRYLNGDKNAYAFTEKHIRIPCFGGEHEVTFIF
                .YRSDVWFVDHPSVHRPNGLYGDKEGAFGNQFRITLLCYAACAPLVLGLGYI
                GONCMFVMDHSLVPLVLLAAKRYPIGYKDSRSLVILHNLAHQGEVAPSTYDPLGL
                PPEWGLAEWPEPNARRHALDGEAVNFKGAVTADRVTVSKGYSWEVTTAEQGO
                GINELLSRKSVLINGVINGIDINDNPNATDKICPHYSVDLLSGKACKGALQKELGL
                PRDPVLPIGIFGRDYGIDLIQILITPDLREDYQVFMLGSGDPELEDNMRKTESI
                FKDKPRGWFVFSVSHRTIAGCDILLMPSEPCGLNQLYAMQIGTVVPVHATGGLR
                FTVENFPFGENGEQGTGWAFPLTENMLTLRTAISTYREHKSSWELMKRGMKSKD
                FTWDAHAQEYQIFOWAFIDRPYVM"
BASE COUNT      582 a      546 c      633 g      622 t
ORIGIN
Query Match      50.3%; Score 1393; DB 6; Length 2383;
Best Local Similarity 83.5%; Pred No. 2.5e-178;
Matches 1619; Conservative 0; Mismatches 310; Indels 10; Gaps 3;

Qy 302 ccgggtgctcgcccccgagtcgcctcgcgccgagtcggcgagcgagccgcggcggt 361
Db 39 CCTCGGCGTGGCGCGCGCGTCCCTCCTCCTCGCGCGCGCGCGCGCGCGCGCGCGT 3;

Qy 362 ccgctcgctcgctcgcgcgcggtccggcgcttggcgggcgctacgctgcgcgagc 421
Db 99 CGGACCGGGCGCGCGCGCGCGGTCTCCAGCGGCTGCTGCGCGCGCGCGCGCGCG 158

Qy 422 tcagcaggaggcgcccgcgcgcccgcccgcccgcccgagtcgagcgagcaactgcgcgc 481
Db 159 TGAGCAGGAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 481

Qy 482 tgccagcttctcgcgcccgcccgcccgcccgcccgcccgagtcgagcgagcgagcgc 218
Db 219 TGCCCGCGCTTCCTCG-----CGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 541

Qy 542 cgccccctgccgagcgccggcggtggggaaactcgccccgacactcctgctcgaaaggattg 601
Db 276 CGCCCGTGCCGACGACGGCGCTGGGGGTCTCGTGTGCGA-----ACCTGAGAGGATTG 329

Qy 602 ctgaggattccatcgacagcataattgtgactgcacagtacacagcagcagcagcagc

```

25



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2002, 12:57:36 ; Search time 249.65 Seconds  
(without alignments)  
9515.915 Million cell updates/sec

Title: US-09-674-824-1

Perfect score: 2771

Sequence: 1 cgcactcactgccttgc.....aaaaaaaaaaaaaaaaaaaaa 2771

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_1101.\*

1:	/SIDS1/gcgdata/geneseq/geneseq/NA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/geneseq/NA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseq/NA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseq/NA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseq/NA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseq/NA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseq/NA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseq/NA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseq/NA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseq/NA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseq/NA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseq/NA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseq/NA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseq/NA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseq/NA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseq/NA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2771	100.0	2805	21	AAZ24487
2	2541.8	91.7	2662	20	AAZ24487
3	2239	80.8	2239	19	AAV01527
4	1393	50.3	2383	19	AAAT95785
5	1392.6	50.3	2491	21	AAZ50636
6	1391.4	50.2	2008	21	AAZ50643
7	1332	48.1	2990	19	AAV66832
8	1329.4	48.0	2533	15	AAQ45183
9	1327.6	47.9	2992	18	AAAT67285
10	1324	47.8	1752	19	AAV29756
11	1322.4	47.7	1749	19	AAV70960

12	1248.4	45.1	1620	19	AAV29759	Zea mays pEXS52 st
13	1215.8	43.9	1528	21	AAZ50637	Corn soluble starch
14	1015	36.6	1415	21	AAZ50642	Corn soluble starch
15	648.4	23.4	1758	17	AAT32325	Soluble starch syn
16	570.8	20.6	1984	21	AAZ46797	Arabidopsis thalia
17	567.6	20.5	10336	20	AAZ34652	Wheat starch solub
18	349	12.6	2380	19	AAV66834	Zea mays soluble s
19	348	12.6	2478	18	AAV67287	Soluble starch syn
20	299	10.8	5072	20	AAZ34653	Wheat starch solub
21	233	8.4	2097	19	AAV29755	Zea mays soluble s
22	233	8.4	2423	19	AAV70958	DNA encoding maize
23	232	8.4	1798	21	AAZ50647	Corn starch syntha
24	232	8.4	2019	21	AAZ50646	Corn starch syntha
25	232	8.4	2248	21	AAZ50651	Corn starch syntha
26	208.6	7.5	2939	21	AAZ86410	Wheat starch synth
27	208.6	7.5	2939	21	AAZ86433	Wheat starch synth
28	195.8	7.1	2807	21	AAZ86435	Wheat starch synth
29	195.8	7.1	2842	21	AAZ86411	Wheat starch synth
30	193.2	7.0	2826	19	AAV01528	Wheat granule-boun
31	190.8	6.9	2007	19	AAV70959	DNA encoding maize
32	190.8	6.9	2007	19	AAV29754	Zea mays soluble s
33	190.8	6.9	2085	18	AAZ67286	Soluble starch syn
34	190.8	6.9	2085	19	AAV66833	Zea mays soluble s
35	186.8	6.7	2107	21	AAZ86412	Wheat starch synth
36	186.8	6.7	2107	21	AAZ86434	Wheat starch synth
37	172.4	6.2	9024	21	AAZ86431	Wheat SSII gene SE
38	155.8	5.6	11384	19	AAV52178	Streptococcus pneu
39	146.6	5.3	2571	19	AAZ96227	S. pneumoniae deri
40	137.6	5.0	2115	21	AAZ32824	Arabidopsis thalia
41	109.8	4.0	2542	19	AAV29753	Oryza sativa waxy
42	104.4	3.8	1275	22	AAZ00044	Wheat cDNA encodin
43	103.4	3.7	2652	18	AAZ75137	DNA encoding a 2.
44	102	3.7	396	22	AAZ94862	Human ovarian canc
45	100.2	3.6	612	22	AAH71471	Human cervical can

#### ALIGNMENTS

RESULT 1

AAZ24487

ID AAZ24487 standard; DNA; 2805 BP.

XX

AC AAZ24487;

XX

DT 18-FEB-2000 (first entry)

XX

DE Wheat soluble starch synthase DNA.

XX

XX Soluble; starch synthase; wheat; transgenic plant; starch production;

KW food; baking; pastry; packaging material; glucose; glucan; paper; pulp;

KW adhesive; textile; building material; soil stabilizer; wetting agent;

KW fertilizer; plant-protection; cosmetic; flocculant; ss.

XX Triticum aestivum.

OS

XX

FH Key Location/Qualifiers

FT CDS 314..2584

FT /\*tag= a

FT /product= "soluble starch synthase"

XX

PN DE19820607-A1.

XX

PD 11-NOV-1999.

XX

PF 08-MAY-1998; 98DE-1020607.

XX

PR 08-MAY-1998; 98DE-1020607.

XX

PA (AGRE ) HOECHST-SCHERING AGREVO GMBH.

XX

PI Loerz H, Luetticke S, Block M;

XX



WPI; 2000-024508/03.  
P-PSDB: AAY50818.

P-PSDB: AAY50818.

New enzyme with starch synthase activity, useful for producing starch for foods and packaging materials -

for foods and packaging materials -

100

Claim 1b; Page 15-19; 24pp; German.

This invention describes a novel protein (I) with the activity of wheat starch synthase. Transgenic plants, specifically wheat, that contain (I) are used for production of starch, used particularly in foods, particularly baked and pastry goods and for making packaging materials or disposable items. Starch may also be used as starting materials for glucose or glucan components (e.g. for fermentation or further chemical conversion); in paper and pulp production, as adhesives, in textiles, in preparation of gypsum-based building materials as soil stabilizer, as wetting agent etc. in fertilizer and plant-protection compositions, as binder (in pharmaceuticals, cosmetics, coal briquetting and casting sand), as flocculant in soil or coal slurries, as rubber and leather additives, and for production of synthetic polymers, e.g. polyurethane films. Transgenic plants with increased/decreased production of (I) produce starches with altered physical and/or chemical properties such as amylose/amylopectin ratios, degree of branching, mean chain length, phosphate content, gelatinization properties, gel- or film-forming properties, or starch grain size or structure. This sequence encodes the soluble starch synthase isolated from wheat (*Triticum aestivum* L. cv. Florida).

sequence 2805 BP; 683 A; 703 C; 763 G; 656 T; 0 other;

```
Query Match          100.0%; Score 2771; DB 21; Length 2805;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2771; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	cgcaactccaactcgcttgcccacactccaactctctctcccgcgcacacgcgaatcgcg	60
Db	35	cgcaactccaactcgcttgcccacactccaactctctctcccgcgcacacgcgaatcgcg	94
Qy	61	acgggtcatcacccatcaactcgctcgccacgcggcaaaccccccgatccgcttttg	120
Db	95	acgggtcatcacccatcaactcgctcgccacgcggcaaaccccccgatccgcttttg	154
Qy	121	caggcagcgcaactaaacccccggggagcgcccccggcgacgacgacgcagtcagtggg	180
Db	155	caggcagcgcaactaaacccccggggagcgcccccggcgacgacgacgcagtcagtggg	214
Qy	181	agagagaggcttcgccccgcgcacccagcggggcgatccacgcctcgctcgctccgca	240
Db	215	agagagaggcttcgccccgcgcacccagcggggcgatccacgcctcgctcgctccgca	274
Qy	241	ctctcgcgctctctccctgtcccgcgcgcacacccatggcggcgacgggctcgcc	300
Db	275	ctctcgcgctctctccctgtcccgcgcgcacacccatggcggcgacgggctcgcc	334
Qy	301	gcggggtgctctgcccccaagctccgcctcgcgccgacatccggcagcgcgcccgggcg	360
Db	335	gcggggtgctctgcccccaagctccgcctcgcgccgacatccggcagcgcgcccgggcg	394
Qy	361	tcgcgcttgctgttcgcgcgcgggctccggcgcttgggcgggcgcgctacgtcgccgag	420
Db	395	tcgcgcttgctgttcgcgcgcgggctccggcgcttgggcgggcgcgctacgtcgccgag	454
Qy	421	ctcagcaggaggccccgcgcgcgcgcccgcgacgacgacaaactggccccgcgcctc	480
Db	455	ctcagcaggaggccccgcgcgcgcgcccgcgacgacgacaaactggccccgcgcctc	514
Qy	481	gtgccaggctctctcgcgcgcgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg	540
Db	515	gtgccaggctctctcgcgcgcgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg	574
Qy	541	ccgccccctgcgcgacgcggcgctggggggaactgcgcgcgcgcgcgcgcgcgcgcgcgc	600



QY 1681 ggcatgatctcattataaaggccattccagagctcatgaggagagcgtgcaatttgc 1740  
 |||||  
 Db 1715 ggcatgatctcattataaaggccattccagagctcatgaggagagcgtgcaatttgc 1774  
 QY 1741 atgcttgatctgggataccaatttttgaagcgtgatgagatctaccagtcagttac 1800  
 |||||  
 Db 1775 atgcttgatctgggataccaatttttgaagcgtgatgagatctaccagtcagttac 1834  
 QY 1801 aagataaattccgtgagtggttgatttagttccagttccacagaaataactgca 1860  
 |||||  
 Db 1835 aagataaattccgtgagtggttgatttagttccagttccacagaaataactgca 1894  
 QY 1861 ggttcgataatgttaatgcctcagatattgaaccttgccgtcttaatacgtatat 1920  
 |||||  
 Db 1895 ggttcgataatgttaatgcctcagatattgaaccttgccgtcttaatacgtatat 1954  
 QY 1921 gctatgaatatggtcagttcctgtagttcatggaactggggccctccgagacacagtc 1980  
 |||||  
 Db 1955 gctatgaatatggtcagttcctgtagttcatggaactggggccctccgagacacagtc 2014  
 QY 1981 gagacttcaaccttttgggtgcaaaagagagagaggttacagggtggcggttccacgc 2040  
 |||||  
 Db 2015 gagacttcaaccttttgggtgcaaaagagagagaggttacagggtggcggttccacgc 2074  
 QY 2041 ctaacgtggacaagatgttgggtgcatgtgcgaacccgcatgtccacattcaggagcac 2100  
 |||||  
 Db 2075 ctaacgtggacaagatgttgggtgcatgtgcgaacccgcatgtccacattcaggagcac 2134  
 QY 2101 aagccctctcggagggttcattgaagcagagcagatgcacaaagaccatcgtgggacct 2160  
 |||||  
 Db 2135 aagccctctcggagggttcattgaagcagagcagatgcacaaagaccatcgtgggacct 2194  
 QY 2161 gccccagcagatcagagcagatcttcagtggtggttccttcgtgacacacccatcagtc 2220  
 |||||  
 Db 2195 gccccagcagatcagagcagatcttcagtggtggttccttcgtgacacacccatcagtc 2254  
 QY 2221 agacgggactggagggttcacagtcgagtccttcctcagtcgaagacatcctctca 2280  
 |||||  
 Db 2255 agacgggactggagggttcacagtcgagtccttcctcagtcgaagacatcctctca 2314  
 QY 2281 tccttcggggccgggaagataccctctacattgctgttctcgtacagtagagtcg 2340  
 |||||  
 Db 2315 tccttcggggccgggaagataccctctacattgctgttctcgtacagtagagtcg 2374  
 QY 2341 caatgcgctgtgttgggttcgggttcgggttcgagacatatgaacggtgctgtgcg 2400  
 |||||  
 Db 2375 caatgcgctgtgttgggttcgggttcgggttcgagacatatgaacggtgctgtgcg 2434  
 QY 2401 gcggtgacagcttcgggtgacacagcttacagtttttggggaataaggagagatgct 2460  
 |||||  
 Db 2435 gcggtgacagcttcgggtgacacagcttacagtttttggggaataaggagagatgct 2494  
 QY 2461 gcaggtatggttaacagaaagcaccactcagatggcagcctctctcgtcgtgttacagct 2520  
 |||||  
 Db 2495 gcaggtatggttaacagaaagcaccactcagatggcagcctctctcgtcgtgttacagct 2554  
 QY 2521 gaaatcagaacacacactggtgactcttttagccttagttagtggaagttgttgcctct 2580  
 |||||  
 Db 2555 gaaatcagaacacacactggtgactcttttagccttagttagtggaagttgttgcctct 2614  
 QY 2581 gttatgttgccttgccttagctgacaaatatttgacctgttggagaattttatcttg 2640  
 |||||  
 Db 2615 gttatgttgccttgccttagctgacaaatatttgacctgttggagaattttatcttg 2674  
 QY 2641 ctgctgtttttttaaatacaaaagaggggtttcctccgatttcattataaaaaa 2700  
 |||||  
 Db 2675 ctgctgtttttttaaatacaaaagaggggtttcctccgatttcattataaaaaa 2734  
 QY 2701 aaaaaa 2760  
 |||||  
 Db 2735 aaaaaa 2794  
 |||||

QY 2761 aaaaaaaaaa 2771  
 |||||  
 Db 2795 aaaaaaaaaa 2805  
 |||||  
 RESULT 2  
 AAX34651  
 ID AAX34651 standard; cDNA; 2662 BP.  
 XX  
 AC AAX34651;  
 XX  
 DT 05-JUL-1999 (first entry)  
 XX  
 cDNA sequence of wheat starch soluble synthase I (SSS I).  
 XX  
 Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBSS;  
 KW starch branching enzyme; starch soluble synthase; debranching enzyme;  
 KW endosperm; wheat; barley; granule-bound synthase; glutenin; starch;  
 KW grain softness protein I; bacterial isoamylase; glycogen synthase;  
 KW wSBE I-D4 gene; ss.  
 XX  
 OS Triticum tauschii.  
 XX  
 PN WO9114314-A1.  
 XX  
 PD 25-MAR-1999.  
 XX  
 PF 11-SEP-1998; 98WO-AU00743.  
 XX  
 PR 20-MAR-1998; 98AU-0002509.  
 PR 12-SEP-1997; 97AU-0009108.  
 XX  
 (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
 PA (GOOD-) GOODMAN FIELDER LTD.  
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.  
 PA (AUSU ) UNIV AUSTRALIAN NAT.  
 XX  
 PI Li Z, Morell M, Rahman S;  
 XX  
 DR WPI; 1999-229525/19.  
 XX  
 PT New isolated cereal plant enzyme genes used for, e.g. expression of  
 PT antisense sequences of granule bound synthase  
 XX  
 PS Claim 10; Page 81-83; 171pp; English.  
 XX  
 CC The invention relates to a novel enzyme of starch biosynthetic pathway  
 CC in a cereal plant, where the enzyme is selected from starch branching  
 CC enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching  
 CC enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or  
 CC SBE I of rice or maize. The methods and products can be used for  
 CC targeting expression specifically to the endosperm of the seeds of cereal  
 CC plants such as wheat or barley. They can be used for the expression of  
 CC e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low  
 CC mol. wt. glutenin, grain softness protein I, bacterial isoamylase,  
 CC bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They  
 CC can be used for modifying the characteristics of starch produced by a  
 CC plant. The present sequence represents the wheat SSS I cDNA sequence.  
 XX  
 SQ Sequence 2662 BP; 592 A; 677 C; 754 G; 639 T; 0 other;

Query Match 91.7%; Score 2541.8; DB 20; Length 2662;  
 Best Local Similarity 98.6%; Pred. No. 0;  
 Matches 2597; Conservative 0; Mismatches 32; Indels 6; Gaps 3;  
 QY 25 ctcccactcttctctcccgcgcacacccagtcgagtcgagccggtcctcaccatcacctcg 84  
 |||||  
 Db 2 ctcccactcttctctcccgcgcacacccagtcgagtcgagccggtcctcaccatcacctcg 61  
 |||||  
 QY 85 gctctgcccaccggaacccccccgatacctcttttcgagcagcagcactaaaccccg 144  
 |||||  
 Db 62 gctctgcccaccggaacccccccgatacctcttttcgagcagcagcactaaaccccg 121  
 |||||







for production of transgenic plants, particularly starch-producing plants, specifically wheat. Use of the isolated nucleic acids, or of antisense sequences, allows starch metabolism to be regulated in transgenic plants. Overexpression may result in improved crop yield, while modification of starch in planta may eliminate the need for subsequent chemical/physical modification. Plants with altered levels of the various isoforms of starch synthase will produce starch of different chain length, amylose/amylopectin ratio, degree of branching, phosphate content, gelatinisation behaviour, granule size and shape, viscosity etc. The starch produced by such plants is useful particularly in foods or to produce packaging materials or disposable goods, as well as in any other known use of starch.

ccccc 2239 BP: 611 A: 448 C: 590 G: 590 T: 0 other;

Query Match	80.8%	Score 2239;	DB 19;	Length 2239;
Best Local Similarity	100.0%	Pred. NO. 1.7e-311;		
Matches 2239;	Conservative	0;	Mismatches 0;	Indels 0;

[illegible][illegible]

b	2282	ccctggtacattgctgtctctctacagttagatgcgaatgcctgcttcg--ttggtc	2340
y	2364	cgccggttcgagaaacatagcagcgtctgtctgcgcgcgcgtgcagccttcgggtggacg	2423
b	2341	cgccggttcgagagtagatgcagcgtctgtctgcgcgcgcgtgcagccttcgggtggatg	2400
y	2424	acagttacagtttggggaataaggaaggatgtctgcaggatggttaacagcaaacga	2483
b	2401	acagttacagtttggggaataaggaaggatgtctgcaggatggttaacagcaaacga	2460
y	2484	ccactcagatgggcagcctctctctgcggtttacagctgaatacagaaaccaactggtgac	2543
b	2461	ccactcagatgggcagcctctctctgcggtttacagctgaatacagaaaccaactggtgac	2520
y	2544	tctttagccttagtgattgtagaagtttgcctctctgtgtatgtctgttgccttagc	2603
b	2521	tctttagccttagcgattgtagaagtttgcctctctgtgtatgtctgttgccttagc	2580
y	2604	tgacaaatatctgacctgttggagaac-----tttatctcttgcctgctgttttttt	2654
y	2601	tatttatttgcctgtttgagaataattttattttatttttgcctgctatttttt	2635

### RESULTS

RESOL 3  
AAV01527  
ID AAV01527 standard: cDNA to mRNA: 2239 BP.

AAV01527;

21-MAY-1998 (first entry)

wheat soluble starch synthase partial-cDNA sequence.

xx starch synthase, wheat: transgenic plant: ss:

XX  
OC  
xxix + i cum aedictum l. cv Florida.

XX	Location/Qualifiers
EW	

FT	Key	DocuID
FT	CDS	3..2018

```
ET
ET
/*tag= a
```

XX PN WO9745545-A1.

04-DEC-1997:

28-MAY-1997: 97WO-EP02793.

XX  
PP  
11-SEP-1996: 96DE-1036917.

PR 29-MAY-1996; 96DE-1021588.

PA (AGRE) HOECHST-SCHERING AGREVO GMBH.

Block M. Loerz H, Luetticke S, Froberg C, Kossmann J;  
XX PT

PI Walter L;

WPI: 1998-032652/03.

DR P-PSDB; AAW23937.

XX PT Nucleic acid encoding starch synthase enzymes from wheat - for  
PT PT transgenic plants that produce modified forms of starch, useful e.g.  
PT PT in foods, or for production of packaging materials and disposable  
PT PT goods

Claim 1: page 47-51; 71pp; English.

CC This near full-length cDNA clone, designated TA55, codes for a  
CC soluble starch synthase (see AAW23837) of summer wheat (cv. Florida).  
CC It was isolated from a phage cDNA library of 21-day-old wheat  
CC caryopses by screening with a PCR fragment derived from rice soluble  
CC starch synthase (see also AAV01529-30). A second clone (see AAV01528),  
CC coding for wheat granule-bound starch synthase (see AAW23938) is also  
CC claimed. These isolated nucleic acids can be inserted into vectors



















Db 543 atagaaattatgcaatttaccagaaaacacattcgcgattccattgcttggcg 602  
Qy 902 gatcacatgaagtaccccttttccatgatatagacaaacgtcgattgggtttgttcg 961  
Db 603 gtgaacatgaagtaccccttccatgatatagatgacgttcgactgggttgggtg 662  
Qy 962 atcatcgttcataacacagaccaggaagtattatgagataaattttgggtcttttggg 1021  
Db 663 atcatcctcatatcacagacctggaaatttatatggagataaagtgttgggttgggtg 722  
Qy 1022 ataactggttcagatacacacactcttgcctgctgctgctgctgctgctgctgctgctg 1081  
Db 723 ataactggttcagatacacacactcttgcctgctgctgctgctgctgctgctgctgctg 782  
Qy 1082 aattgggaggatatattatggacagaaattgcatgtttgttggtaacgattggcatgcc 1141  
Db 783 aattgggaggatatattatggacagaaattgcatgtttgttggtaacgattggcatgcc 842  
Qy 1142 gcttggcagtcctcttctgctgcaaaatatagaccatacgggttttaccagagattccc 1201  
Db 843 gtctagtgccagtcctcttctgctgcaaaatatagaccatacgggttttaccagagattccc 902  
Qy 1202 gcagcaccccttgttatacataatttagcacatcaggtgtgagcctgcagatacatc 1261  
Db 903 gcagcattctgttaatacataatttagcacatcaggtgtgagcctgcagatacatc 962  
Qy 1262 ctgatactggagtcctctgctgcaaaatatagaccatacgggttttaccagagattccc 1321  
Db 963 ctgacctgggttggccacctgaaatggtatggagctcgtgagtggtattccctgaaatggg 1022  
Qy 1322 caaggagcagtcctctgacaaaggtgagcagcttaacttttgaaggagcaggtgtgta 1381  
Db 1023 cgaggagcagtcctctgacaaaggtgagcagcttaacttttgaaggagcaggtgtgta 1082  
Qy 1382 cagcagatcgattgtgacgcctgacgcaggttattcattgagggaggttcacaaactgctgaag 1441  
Db 1083 cagcagatcgattgtgacgcctgacgcaggttattcattgagggaggttcacaaactgctgaag 1142  
Qy 1442 gtggacagggccctcaatgagctcttaagctcccgaaaaagtgtattgaaatggaaatgtaa 1501  
Db 1143 gtggacagggccctcaatgagctcttaagctcccgaaaaagtgtattgaaatggaaatgtaa 1202  
Qy 1502 atggaattgacataatgattgaaacccacacacaggtgtctccctcattcattatt 1561  
Db 1203 atggaattgacataatgattgaaacccacacacaggtgtctccctcattcattatt 1262  
Qy 1562 ctgctgatgacctctctggaaggccaaatgtaaaagctgaattgcagaaagagttgggtt 1621  
Db 1263 ctgctgatgacctctctggaaggccaaatgtaaaagctgaattgcagaaagagttgggtt 1322  
Qy 1622 taactgtaaggaggaggtgttccctcctgattggtcttatttgaagagctggattaccagaaag 1681  
Db 1323 taactgtaaggaggaggtgttccctcctgattggtcttatttgaagagctggattaccagaaag 1382  
Qy 1682 gcatgattcattcaaaatggccattccagagctcattgagggagagcgtgcaatttgctca 1741  
Db 1383 gcaattgatccattcaaaatggccattccagagctcattgagggagagcgtgcaatttgctca 1442  
Qy 1742 tgcttgatctggggggtcccaatttttgaaggctggatgagatcaccgagctgaggtaca 1801  
Db 1443 tgcttgatctggggggtcccaatttttgaaggctggatgagatcaccgagctgaggtaca 1502  
Qy 1802 aggaataattccgtgagatgggttggatttagtgggttccagtttcccccagaaataaactgcag 1861  
Db 1503 aggaataattccgtgagatgggttggatttagtgggttccagtttcccccagaaataaactgcag 1562  
Qy 1862 gttgcgatatattgttaattgcatcagatttgaaccttgcggtcttatacagctatatg 1921  
Db 1563 gctgcgatatattgttaa tgcgacccagattcgaaccttgcggtcttatacagctatatg 1622  
Qy 1922 ctatgcaatatggtacagttcctgtagttcattggaactgggggctccgagacacagctcg 1981

Db 1623 ctatgcagtatggcacagttcctgttgcctatgcacatgggggcttagagataccgtgg 1682  
Qy 1982 agaccttcaacccttttgggtgcaaaaaggagaggttacagggtgggttctcaccgc 2041  
Db 1683 agaacttcaacccttttgggtgcaaaaaggagaggttacagggtgggttctcaccgc 1742  
Qy 2042 taaccgtggacaagatgttgggtgcaattgcgaaccgcgattgcgacattccaggagagcaca 2101  
Db 1743 taaccacagaaaacatgttgggacattgcgaactgcgaatctacatacagggaaacaca 1802  
Qy 2102 agcgtctcctggggggggtcctcatgaagcgagcgcattgacgaagaccatacgtgggaccatg 2161  
Db 1803 agtctcctgggaagggttaataagcgaggtcatgtcaaaagacttcaactcaggtgggaccatg 1862  
Qy 2162 cc-ccgagcagtcagcagacagatcttgcagtggtggccttcgtggaccaaccctacgtcatgt 2220  
Db 1863 ccgctgacaatacgaacaaattctccagttggccttcacgtcgatcgaccctatgtcatgt 1922  
Qy 2221 aga 2223  
Db 1923 aaa 1925

RESULT 7  
AAV66832  
ID AAV66832 standard; DNA; 2990 BP.  
XX AAV66832;  
XX AC  
XX AAV66832;  
DT 05-JAN-1999 (first entry)  
XX  
DE Zea mays soluble starch synthase gene SSS1052 and SSS64.  
XX  
KW Zea mays; US yellow-dent corn line; maize; soluble starch synthase;  
KW SSS; glycogen biosynthetic pathway; branching enzyme; ss.  
XX Zea mays.  
XX OS  
XX US5824790-A.  
XX PN  
XX PD 20-OCT-1998.  
XX  
XX PF 15-DEC-1995; 95US-0572951.  
XX PR 15-DEC-1995; 95US-0572951.  
PR 21-JUN-1994; 94US-0263921.  
PR 29-NOV-1994; 94US-0346602.  
XX  
XX PA (ZENE ) ZENECA LTD.  
XX  
PI Guan H, Keeling PL, Knight MB;  
XX  
DR WPI; 1998-582626/49.  
XX  
PT Isolated nucleic acid molecule, used to produce transgenic plants -  
PT comprises nucleotide sequence encoding polypeptide having soluble  
PT starch synthase activity, where polypeptide is encoded by maize gene  
XX  
PS Claim 1; Column 25-28; 29pp; English.  
XX  
CC The present sequence represents an isolated nucleic acid molecule which  
CC has been isolated and comprises a nucleotide sequence encoding a  
CC polypeptide having soluble starch synthase (SSS) activity, where the  
CC polypeptide is encoded by a maize gene. The isolated nucleic acid  
CC molecule can be used to produce transgenic plants with altered starch  
CC production. The transgenic plants produced using the nucleic acid  
CC molecule have an enhanced ability to produce structurally-altered starch.  
XX  
SQ Sequence 2990 BP; 758 A; 655 C; 801 G; 776 T; 0 other;

Query Match 48.1%; Score 1332; DB 19; Length 2990;  
Best Local Similarity 81.7%; Pred. NO. 4.7e-182;







transition; protoplast; expression; ss.

*Oryza sativa.*

**Key** Location/Qualifiers

```
3 01K
T::TTS
/*tag=
```

```

CDS
II4:::IS
/*tag=

```

sig\_peptide

mat\_peptide
453...19
/\*trac=

3'UTR  
1992...2  
/4499-

0770700

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

[illegible]

.....

P-PSDB; AAR51231.

Soluble rice starch syn

amyloplast

Claim 1: page 15-17: 18

The soluble rice starch

the transit peptide that  
enzymes to the amyloid

protoplast augments the

the efficient transition

Sequence 2533 BP; 630 A

### Local Similarity

100

[illegible][illegible]

24

TTT

၁၂

I/4 CCygggagggcgcgc

430 gagggcccccgcg

234 gacggTgggTcgggcgcac

484 ccaggcttcctcgcgccg

294 ccgaccttcctcgtgccg

538 cagcgccctgccgga

354 ccgaccccgccgcttq

KW	transliteration; protoplast; expression; ss.
XX	
OS	Oryza sativa.
XX	
PH	Location/Qualifiers
FT	1..113
FT	/*tag= a
FT	114..1994
CDS	/*tag= b
FT	114..452
FT	sig_peptide
FT	/*tag= c
FT	453..1991
FT	mat_peptide
FT	/*tag= d
FT	1992..2533
FT	/*tag= e
FT	
XX	
PN	JP06070779-A.
XX	
PD	15-MAR-1994.
XX	
PF	07-JUL-1992; 92JP-0179947.
XX	
PR	07-JUL-1992; 92JP-0179947.
XX	
PA	(MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
XX	
WEI	1994-128678/16.
DR	P-PSDB; AARS1231.
XX	
PT	Soluble rice starch synthetic enzyme gene and transit peptide -
PT	for the efficient transport of heterologous proteins to
PT	amylolplast
XX	
PS	ClaIm 1; Page 15-17; 18pp; Japanese.
XX	
CC	The soluble rice starch synthetic enzyme gene has, at the N-terminal,
CC	the transit peptide that is required for the transition of this
CC	enzyme to the amyloplast. Introduction of this gene into the rice
CC	protoplast augments the expression of soluble rice starch synthetic
CC	enzyme. The transit peptide coding sequence can be used for
CC	the efficient transition of any protein into amyloplasts.
XX	
SQ	Sequence 2533 BP; 630 A; 550 C; 701 G; 652 T; 0 other;

Query Match 48.0%; Score 1329.4; DB 15; Length 2533;  
Best Local Similarity 80.7%; Pred. No. 1.1e-181;  
Matches 1606; Conservative 0; Mismatches 371; Indels 14; Gaps

[illegible]



Db 1494 aaaggcatgatcataataaattgcattccagatctcatcgcgacaaattcaattc 1553  
Qy 1738 gtcatcttgatctggggaatccaatttttgaaggctgagatgagatctaccgagtgagt 1797  
Db 1554 gtcatcttgatctggggaatccaatttttgaaggctgagatgagatctaccgagtgagt 1613  
Qy 1798 tacaagataaaatccgttgatgggtgattagttgattccagtttccacagaataaact 1857  
Db 1614 tacaggataaaatccgttgatgggtgattagttgattccagtttccacagaataaact 1673  
Qy 1858 gaaggttcgatatattgttaatgcacatgcagagattgaaacctgctgttttaatacagcta 1917  
Db 1674 gcaggttcgatatattgttgatgcacatgcagagattgaaacctgctgttttaatacagcta 1733  
Qy 1918 tatgctatcaaatgatgtacagtctctgatttctagttcatggaactggggcctccagagaca 1977  
Db 1734 tatgctatgcaaatgatgtacagtctctgatttctagttcatggaactggggcctccagagata 1793  
Qy 1978 gtgagacattcaacctttttgtgcacaaaggagagagaggttacaggtgggcttctca 2037  
Db 1794 gtgagaattttaacctgttctgagaaaggagagagaggttacaggtgggcttctca 1853  
Qy 2038 ccgctaacc-gtggacaaagattgttggtcattgcgaacgcgagatgacatcagggga 2096  
Db 1854 ccaataaccattgaaacaaatgctgtggcattgcgagtggaatttcgacatacagggga 1913  
Qy 2097 gcacaagcgtctctggaggggtctatgaagcgagagcatgacgaagaccatcgtggga 2156  
Db 1914 acacaagtcctcttgaggaggttctaataagcgagagcatgctcaagcgactttaca 1973  
Qy 2157 ccattgcc-ccgagcagctacagcagatctctcagttgggcttctgtagaccacccact 2215  
Db 1974 ccattgccctcacagatgaaacagatcttcgaatggccttcattgatacaacatatgt 2033  
Qy 2216 catgtagacgg 2226  
Db 2034 catgtaaatgg 2044

RESULT 9  
AAT67285  
ID AAT67285 standard; cDNA; 2992 BP.  
XX  
AC AAT67285;  
XX  
DT 11-SEP-1997 (first entry)  
XX  
DE Soluble starch synthase cDNA clone SSS10.52.  
XX  
KW Soluble starch synthase; SSS10.52; transgenic plant; cereal; maize;  
KW ss.  
XX  
OS Zea mays inbred line W64A.  
XX  
PN WO9720936-Al.  
XX  
PD 12-JUN-1997.  
XX  
PF 04-DEC-1996; 96WO-GB02990.  
XX  
PR 06-DEC-1995; 95GB-0024938.  
XX (ZENE ) ZENECA LTD.  
XX  
PI Keeling PL, Knight ME;  
XX  
DR WPI: 1997-319782/29.  
XX  
PT cDNA encoding soluble starch synthase - used to produce transgenic  
XX plants with increased capacity for producing and storing starch  
PS Claim 5; Page 17-20; 44pp; English.

XX  
CC cDNA clones SSS10.52, SSS6.31 and SSS56 (AAT67285-87) code for maize  
CC soluble starch synthase (SSS). They were isolated from a maize  
CC inbred line W64A library by screening with a probe (AAT67288) based  
CC on pea SSS, rice SSS and maize GBSS. The cDNAs can be used to  
CC produce plants with an increased capacity for producing starch, or  
CC a capacity to produce starch with an altered fine structure. They  
CC can also be used to isolate the corresponding genomic sequences  
CC from crop plants, to determine the contribution of the SSS gene to  
CC the net regulation of starch biosynthesis, and to modify the levels  
CC of starch produced by the plant. Transgenic plants (esp. maize)  
CC can be used to produce hybrid plants which have higher rates of  
CC starch synthesis at temperatures above the normal optimum.  
XX  
SQ Sequence 2992 BP; 758 A; 655 C; 801 G; 776 T; 2 other;

Query Match 47.9%; Score 1327.6; DB 18; Length 2992;  
Best Local Similarity 82.0%; Pred. NO. 2e-181;  
Matches 1606; Conservative 0; Mismatches 335; Indels 18; Gaps 6;  
Qy 247 cgcctctctctctctctcccgcgccacacacccatggcgcgagggcgctcgccgcccgg 306  
Db 563 ccccccccccgctctctgtgtctctctctctccgaatggcgagcgcctcgccgtg 622  
Qy 307 tgctcgccccacagcgtccgctgcgcgacgacgacgacgacgacgacgacgacgacgac 366  
Db 623 ggcgcgcgtgctctctctctctcgcgggcgctgcgcgcgcgcgcgcgcgcgcgcgcgc 676  
Qy 367 tgctgctcgc 426  
Db 677 cgggcgccgc 736  
Qy 427 agggagggcccg 486  
Db 737 agggagggcccg 796  
Qy 487 ggcttctctcg 546  
Db 797 ggcttctcg--cg 853  
Qy 547 ctgcccgc 606  
Db 854 gtgcccgc 907  
Qy 607 gattccatgcacagcataattgtgctgcaagtgcagagatctcagatcagatgagcg 666  
Db 908 ggtccatgcataacacagtagttgtggcaagtgcagagatctcagatcagatgagcg 967  
Qy 667 aatgagcaacctcaagctaaagtacacgtagcatcggtttgtgactggtgagctgct 726  
Db 968 aaggagcagctcagctaaagttaacaaagcattgctcttgaaccgcgagcttct 1027  
Qy 727 cctt-atgcaaaagtcagggggttgaggagatgtttgtgcttcctaccatattgctctgc 785  
Db 1028 ccttaatcgaaagctggtggtctaggagatgttgggttcattgccaagtgtctctgc 1087  
Qy 786 tgctcgtgtccagagtgatgttgaatgccaagatacctaataaggggtccctcgataa 845  
Db 1088 tgctcgtgtccagctgtgattgttaatgccaagatacctaataaggggtccctcgataa 1147  
Qy 846 aaactatgcaaaaggcattatcacactgcgaagcacataagatccatgcttgggggagtc 905  
Db 1148 gaattatgcaaatgcaatttactacagaacacacattccgattccattcttggcggtga 1207  
Qy 906 acatgaagtgcacctttttcatgagtatagagacaacgcgtggtggtttgttcgatca 965  
Db 1208 acatgaagttaaccttctccatgagtatagagatacattcagttgactgggtgtttgtgatca 1267  
Qy 966 tccgtcatatcacagaccagggaagtattatgagataaatttgggtgcttttgggtgataa 1025  
Db 1268 tccgtcatatcacagaccctggaaatttataggagataaatttgggtgcttttgggtgataa 1327







Qy	1609	aaggagtggtttaa	ctgctgaaggagagatg	tctctctgattgcttattgaa	aactg	1666
Db	1192	aaggagtggtttaa	ctgctgaaggagatg	tctctctgattgcttattgaa	ggttg	1251
Qy	1669	gattacagaaga	gcattgatctcattaaat	ggccattccagagctcatgaggagagac		1728
Db	1252	gattatcagaaga	gcattgatctcattaaat	ggccattccagagctcatgaggagagac		1311
Qy	1729	gtgcaatttgtca	tgcttggattcgggatgc	caatttttgaaggctggatgagatctacc		1788
Db	1312	gttcaaatgttca	tgcttggattcgggatgc	caatttttgaaggctggatgagatctacc		1371
Qy	1789	gagtcagagttaca	agagataaaattcgttggatg	gggttgatttagtittccagtttccccac		1848
Db	1372	gagtcagatctca	agagataaaattcgttggatg	gggttgatttagtittccagtttccccac		1431
Qy	1849	agataaactcag	gttgcgatattattgtaat	gcacatcgagatttgaaccttgcggtctt		1908
Db	1432	cgaataactcgg	gtcgcgatattattgtaat	gcacatcgagatttgaaccttgcggtctt		1491
Qy	1909	aatcagctatat	ctctatgacaatatggtac	agttctctgatttgaaccttgcggtctt		1968
Db	1492	aatcagctatat	ctctatgacaatatggtac	agttctctgatttgaaccttgcggtctt		1551
Qy	1969	cgagacacactc	agacaccttcaacctttt	gtgcataaagagagaggtacaggttg		2028
Db	1552	agagataccctg	gagaaacttcaacctttt	gtgcataaagagagaggtacaggttg		1611
Qy	2029	gcgttctcacgc	gttaaccgttgacaagatg	-ttgtgggcattgcgaaccgcgattgcgac		2087
Db	1612	gcattgcaccc	cttaaccagaaaaatgt	tcttgacattgcgaactgcaatattcac		1671
Qy	2088	attcaggaggaca	ccgctcctctggagggtc	atgaaagcagaggtacaggaagacca		2147
Db	1672	atacagggaaca	agctcctctctggaaagg	cttaataagcagaggtacgaaagactt		1731
Qy	2148	tacgtggaccat	gcccc	2165		
Db	1732	cactggggaccat	gcccc	1749		
RESULT 11						
AAV70960						
ID	AAV70960 standard; DNA; 1749 BP.					
XX						
AC	AAV70960;					
XX						
DT	23-AUG-1999 (first entry)					
XX						
DE	DNA encoding maize starch soluble synthase I-2.					
XX						
KW	Non-glycogen-like polysaccharide production; fermentation;					
KW	starch synthesis enzyme; glycogen-synthesis enzyme; glycogen synthesis					
KW	non-starch branching gene; amylopectin; amylose; plant-like starch;					
KW	maize starch soluble synthase I-2; ss.					
XX						
OS	Zea mays.					
XX						
PN	W09844780-A1.					
XX						
PD	15-OCT-1998.					
XX						
PF	03-APR-1998; 98WO-US06660.					
XX						
PR	04-APR-1997; 97US-0042939.					
XX						
PA	(EXSE-) EXSEED GENETICS LLC.					
XX						
PI	Guan H, Keeling PL;					
XX						
DR	WPI; 1998-568285/48.					
DR	P-PSDB; AAV70894.					



[illegible]

Db	652	cttttgatccttgaaattggaagagataatttatggacagaaattgcatgtttgtgtgccaat	7111
Qy	1129	gattggcaatgccagcccttgccagctcctcttctgtgcgaaataatagaccatacagtggtt	1188
Db	712	gattggcaatgccagctcagtgccagtccttctgtgcgaaataatagaccatacagtggtt	771
Qy	1189	tacagagattcccgagcacccctgttatacataatttagcacatcagggtgtggagcct	1248
Db	772	tataaagactcccgagcatctctgttaaataataatttagcacatcagggtgtgagcct	831
Qy	1249	gcaagtacatatcctgatactgggattgcccctcctgaatgggtatggagctttagaatgggta	1308
Db	832	gcaagcacatacctgacctgggttgcacactgaatgggtatggagcctctggagtggggta	891
Qy	1309	tttcagaaatggcagagaggcatgcccttgacaagggtgagggcagtttaactttttgaaa	1368
Db	892	ttcctgaaatggcgagagaggcatgcccttgacaagggtgagggcagtttaactttttgaaa	951
Qy	1369	ggagcagttatgacagcagatcgattgtgacccgtcagtcagggttatctcatggagagtc	1428
Db	952	ggtagcattgtgacagcagatcgaaatcgtagctgcagttaagggttatctggggagagtc	1011
Qy	1429	acaactgtgaagtgtagcagggcccaatgagctctttaaagctcccgaaaaagtgatttg	1488
Db	1012	acaactgtgaagtgtagcagggcccaatgagctctttaaagctcccgaaaaagtgatttg	1071
Qy	1489	aatggaattgtaaatggaattgacatgaattgattgggaaccccaacacagacaagtgcttc	1548
Db	1072	aacggaattgtaaatggaattgacatgaattgattgggaacccctgcacagacaaatgtatc	1131
Qy	1549	cctcatcatattctgtcgatgacctctctggaaggcccaaatgttaaagctggaattgcag	1608
Db	1132	ccctgtcatattctgttgatgacctctctggaaggcccaaatgttaaagctggaattgcag	1191
Qy	1609	aaggagttgggtttacctgttaagggaggagtgtcctctgattggtctttattggaagactg	1668
Db	1192	aaggagctgggtttacctataaggccctgagtgtcctctgattggtctttattggaagctg	1251
Qy	1669	gattaccagaaaggcattgatctcatataaaatgggccattccagagctcatgagggaggac	1728
Db	1252	gattatcagaaggcattgatctcatcctaactatcataccagatcctatgcggggaagat	1311
Qy	1729	gtgcaattgtcatgtgttgattggatctggggatcccaatttttgaggctggatgagatctacc	1788
Db	1312	gttcaattgtcatgtgttgattggatctgggtgaccagagcttggaagattggatgagatctaca	1371
Qy	1789	gagtcgagttacaaagataaaatccgtgagtggttgattgatttagtcttcagatttccac	1848
Db	1372	gagtcgagttcgaaggataaaattcgtggattgggttgatttaggttccagtttccacac	1431
Qy	1849	agaataactgcaggttcgatataattgttaaagccatcgagattgaaaccttgcggctctt	1908
Db	1432	cgaaataactgcggctcgatataattgttaaagccatccagatcgaaccttgcgtctc	1491
Qy	1909	aatcagctatatgctatgcaatatggtacagttcctctgatttcatgaaactgggggcctc	1968
Db	1492	aatcagctatatgctatgcaatatggtacagttcctctgatttccatgaaactgggggcctt	1551
Qy	1969	cgagacagatcgagaccttcaaaccttttgggtgcgaaaggagaggggtacaggggtg	2028
Db	1552	agagataactggaagaacttcaaaccttttgggtgagaatggagagcagggtacaggggtg	1611
Qy	2029	gcgttctacccgctaaccctggacaagatg-ttggtgggcatctgcgaaccgcgatgtcgac	2087
Db	1612	gcattgcacccctaacccagaaaaaatgtttgtggacattgcgaactgcgaactctac	1671
Qy	2088	attcagggagcacaaagcctctctggaggggctcatgaagcgagcagatgcgaaagacca	2147
Db	1672	atcacgggaacaaagccctctctgggaagggttaatgaagcgagggcatgtcaaaagactt	1731
Qy	2148	tactgtggaccatgcccc	2165
Db	1732	cagctgagaccatgccac	1749







Db 1264 cgtggatgggtggtgatttagttccagtttccacogataaactgcccgtgcgata 1323  
 QY 1873 ttgttaatgccatcgagatttgaaccttgcgtcttaactcagcttatgtatgcaatat 1932  
 Db 1324 ttgttaatgccatcgagatttgaaccttgcgtcttaactcagcttatgtatgcaatat 1383  
 QY 1933 ggtacagtttctgtagttcatggaactggggccctccgagacacacagtcgagacctcaac 1992  
 Db 1384 ggcacagtttctgttccatgcaactggggcccttagagataccgtgggagaacttcaac 1443  
 QY 1993 ccttttggtcacaagagaggggggtacaggggtggggttcttccacgcgttaacccgtggac 2052  
 Db 1444 cctttcggtgagaaatgagagcaggggtacaggggtggggtcattcgaccccttaaccacagaa 1503  
 QY 2053 aagatg-ttgtggcattcgaaacccgagatgctgacattcagggagacacagcgcgtctg 2111  
 Db 1504 aacatgttttgagacttgcaactgcaactgcaacttcaatctacatacagagaaacacagtcctctg 1563  
 QY 2112 ggaaggggtcattgaagcgagggcatgacgaagaccatacgtgggaccatgcccc 2165  
 Db 1564 ggaagggtaatgaagcgagggcatgtcaaaagacttcacgtgggaccatgccc 1617

RESULT 13

AAZ50637/c  
 ID AAZ50637 standard; cDNA; 1528 BP.

XX AAZ50637;  
 AC AAZ50637;  
 XX  
 DT 23-MAY-2000 (first entry)  
 XX  
 DE Corn soluble starch synthase gene fragment inserted in pSS42.  
 XX  
 KW Soluble starch synthase; starch fine structure; corn;  
 KW transgenic plant; amylose; amylopectin; amylose polymerisation;  
 KW non-granule bound starch synthase; non-GBSSI; altered starch; food;  
 KW paper; plastic; adhesive; ss.  
 XX  
 XX 2ea may.

OS WO200006755-A2.  
 PN

XX 10-FEB-2000.  
 PD

XX 26-JUL-1999; 99WO-US16296.  
 PF

XX 28-JUL-1998; 98US-0094436.  
 PR

XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 PA

XX Broglie KE, Lightner JE;  
 PI

XX WPI; 2000-195311/17.  
 DR

XX Producing transgenic cereal crops with altered starch structure useful  
 PT for preparing foodstuff, paper, plastic or adhesives, comprises  
 PT transforming crops with chimeric sense or antisense gene construct  
 PT encoding starch synthase  
 XX

PS Claim 5; Page 51-52; 56pp; English.  
 XX

CC The present sequence is the corn soluble starch synthase (SSI) DNA  
 CC sequence inserted into plasmid pSS42. Chimeric gene of pSS42 was cloned  
 CC into the vector pKS17 for the generation of an antisense construct for  
 CC suppression of SSI expression in corn. The starch fine structure derived  
 CC from a grain of the cereal crop can be altered in the transformed cereal  
 CC crop by changes in amylose to amylopectin ratio, amylopectin fine  
 CC structure, increased abundance of very short amylopectin chains and in  
 CC the degree of polymerisation of amylose. These modifications can be  
 CC created by controlling the expression of non-GBSSI (non-granule bound  
 CC starch synthase) in transgenic plants. Altered starches are useful in  
 CC foods, paper, plastics or adhesives.

XX SQ Sequence 1528 BP; 440 A; 383 C; 300 G; 405 T; 0 other;  
 Query Match 43.9%; Score 1215.8; DB 21; Length 1528;  
 Best Local Similarity 87.7%; Pred. No. 1.8e-165;  
 Matches 1399; Conservative 0; Mismatches 187; Indels 1; Gaps 1;  
 QY 678 tcaagctaaagtacacgttagcatcgtgtgtgtgactgggtgaagctctctcttgcacaa 737  
 Db 1528 TCGAGCTAAAGTAACACAAAGCATGTCTTTGTAAACGGGGAAGCTTCTCTTATGCAAA 1469  
 QY 738 gtacaggggggttggagatgtttgtgttccgtttaccattgctcttgcgtcgtgctca 797  
 Db 1468 GTCTGGGGTCTAGGAGATGTTTGTCTCAATTCACAGTTCCTTCTGCTGCTGCTGCTCA 1409  
 QY 798 ccagtgatggttgaatgccaagatacttaaatgggtccctctgtataaaaaactatgacaa 857  
 Db 1408 CCGTGTGATGGTGTAAATGCCAGATATTTAAATGGTATGCCATTCAGTAAGAAATATGCAAA 1349  
 QY 858 ggcattatacactgcgaagcacattaaagattccatgcttttgggggatacacatgaagtac 917  
 Db 1348 TGCATTTTACACAGAAAAACACATTCGATTCATGCTTTGGCGGTGAACATGAAGTTAC 1289  
 QY 918 ctttttcatgagtatagagacaaacgtcgattgggtgttbtgtgcatacatcogtcataca 977  
 Db 1288 CTTCTTCATGAGTATAGAGATTTCAGTTGACTGGGTGTTTGTGATCATCCCTCATATCA 1229  
 QY 978 cagaccaggaagtttatagagataattttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1037  
 Db 1228 CAGACCTGGGAAATTTATGGAGATAAGTTGGTGGCTTTGGTGATTAATCAGTTTCAGATA 1169  
 QY 1038 cactactcttgcattgctgcagtcgagggcccccactaactccttgaattgggagagatatat 1097  
 Db 1168 CACACTCCTTTGCTATGCTGCATGTGAGGCTCCTTTGATCCTTCAATTTGGGAGGATATAT 1109  
 QY 1098 ttatggacagaattgcatgtttgttgaacgatttggcatgccgcgttbtgtccagtcct 1157  
 Db 1108 TTATGGACAGAAATTCGATGTTTGTGTCATGATGGCATGGCCAGTCTAGTGCCAGTCCCT 1049  
 QY 1158 tcttgcgtcaaaatatagacatacaggtgttttacagagattcccgagacccttcttat 1217  
 Db 1048 TCTTGTCTGCAAAATATAGACCATATGTTGTGTTTATAAGACTCCCGCAGCATCTTGTGAAT 989  
 QY 1218 acataattatgacacatcaggggtgtggagcctgcaagtacatactctgtatcgtggattgcc 1277  
 Db 988 ACATAATTTAGCACATCAGGGGTGTAGAGCTGTGAAGCAGACATATCCTGACCTTGGGTTGCC 929  
 QY 1278 tcttgaattggtatgagcttttagaatgggtatttccagaatgggcaaggagggcatgccc 1337  
 Db 928 ACCTGAATGGTATGGAGCTCTGGAGTGGGTATTCCTTGAATGGGCGAGGAGCATGCCCT 869  
 QY 1338 tgacaaggttgaggcagtttaacttttgaagagagcagttgtgacacagatcgagattgt 1397  
 Db 868 TGACAAGGGTGAGGCAGTTAATTTTGAAGGTGCAAGTTGTGACACAGATCGAATCGT 809  
 QY 1398 gaccgtcagtcaggtttattcatggtggaggttcacaactgctgaagtgagagggcctcaa 1457  
 Db 808 GACTGTCTAGTAAGGGTTATTTCGTGGGAGGTCAACAATGCTGAAGGTGGACAGGGCCCTCAA 749  
 QY 1458 tgagctcttaagctcccccagaaagtgtatgaaatggaaatgtaaatgaaatgacataa 1517  
 Db 748 TGAGCTCTTAAGCTTCCCAAGAGAGTGTATTAACGGAAATTTAAATGGAATTTGACATTA 689  
 QY 1518 tgattgaaacccaccacacagcagttctccctcatcattattctctatcagaccctctc 1577  
 Db 688 TGATTGGAACCCCTGCCACACAGCAAAATGTATCCCTCTGTCAATTTCTGTTGATGACCTCTC 629  
 QY 1578 tggaaagcccaaatgtaagcgtgaattgcagaagaggttgggtttacctgttaaggaggga 1637  
 Db 628 TGGAAAGCCCAAAATGTAAGGTGCTATTCAGAGAGGAGCTGGGTTTACCTATAGGCTCGA 569  
 QY 1638 tgttccctctgattgggtttatttggaaagactggattacacagaaaggaattgatctcataa 1697











QY 1456 aatgagctcttaagctcccgaaaaaagtgtattgaattggaattgaaattgacatt 1515  
Db 610 catgagctgttgagcagtagcagctgtcttcttaattggaattactaatggaatgagctt 669  
QY 1516 aatgattgaaacccaccacagacagaagtgtctccctcatcattattctgtcgcgatgcctc 1575  
Db 670 aatgattgaaacccgtcgacagatgagcatatcgcttcgacttaactccatcaatgacctc 729  
QY 1576 t-----ctggaagaagccaaatgaaactgaaatgcagaagagtggtttacctgta 1629  
Db 730 tccccccctggaaggttcagtgcaagactgactgcaaaaggaactggccttccaatt 789  
QY 1630 agggagagatgtcctctgcttattgtgaagactggaattaccagaaagcattgat 1689  
Db 790 cgacccgattgtccactgattgatttattggaagctggactaccagaaagtggtgac 849  
QY 1690 ctcaataaattgccattccagagctcatgagggagagcgtgcaatttgcattgcttgga 1749  
Db 850 ataactctgtcagcaattccagaacttaigcagaatgagtccaagtgttaatgcttggga 909  
QY 1750 tctggggatccaatttttgaagcctgagatgagatctaccgagtcgaggttacaaaggataaa 1809  
Db 910 tctgggtgagaacaatagaaagactgagagacatacacagaaaatctttttaaagacaa 969  
QY 1810 tccgtggatgggttgatttagttccagttccagttccacagaaataactgcaggttgcgat 1869  
Db 970 ttcggtgcttgggttgatttaattgttccagtttctcattaggaatacagcaggtatcgac 1029  
QY 1870 atattgttaatgccatcgagatttgaaccttgcggtcttaatacagctatatgtctatgcaa 1929  
Db 1030 atactattgatgcctcaagattcgaacctgtggtttaaaccatgttatgcaatgaga 1089  
QY 1930 tatggtacagttcctgttagttgaaactggggggtcccgagacacagtcgagaccttc 1989  
Db 1090 tatggcaccatacctattgttctatagcacggtggggttaagagacacagtgaaaggatttt 1149  
QY 1990 aaccctttgtgcataaagg---agaggaggtacaggttgggttctcaccgcctaacc 2046  
Db 1150 aatccatatgtcgaagaagaaaggtagaaggtaccggttggacattttctcctctaagc 1209  
QY 2047 gtggacaagatgttgtgggcatgcaaccgcgagtgcgacattcaggggagcacaagccg 2106  
Db 1210 agtgaaaagtgttgatacacactgaagctggcgatcaggacttatcacagaacataagtc 1269  
QY 2107 tcttgaggaggtcctatgaagcagagcatgacgaagaaccatacgtgggaccatgc-ccc 2165  
Db 1270 tcttgaggaggttgcataagagaggtatgggaaggagactattcctctgggaaatgcagcc 1329  
QY 2166 gacgagtacgagcagatcttcgagtggtggccttcgtggaccaaacccctacgtca 2217  
Db 1330 attcaatatgagcaagttttcaactgggcttttatagatcctccatattgca 1381

Search completed: March 28, 2002, 15:01:12  
Job time: 7416 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen. Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2002, 12:13:06 ; Search time 104.55 seconds  
(without alignments)  
6002.588 Million cell updates/sec

Title: US-09-674-824-1  
Perfect score: 2771  
Sequence: 1 cgcacactccactgccttgc.....aaaaaaaaaaaaaaaaaaaaa 2771

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues  
Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2239	80.8	2239	4	US-09-196-390-1
2	1393	50.3	2383	4	US-09-192-909-1
3	1332	48.1	2990	1	US-08-572-951-1
4	1325.6	47.8	1752	3	US-08-941-445A-12
5	1248.4	45.1	1620	3	US-08-941-445A-20
6	752.8	27.2	2360	3	US-08-836-567-9
7	648.4	23.4	1758	3	US-08-836-567-3
8	349	12.6	2380	1	US-08-572-951-3
9	287	10.4	2793	3	US-08-836-567-7
10	279	10.1	1926	3	US-08-836-567-5
11	233	8.4	2097	3	US-08-941-445A-10
12	193.2	7.0	2825	4	US-09-196-390-5
13	190.8	6.9	2007	3	US-08-941-445A-8
14	190.8	6.9	2085	1	US-08-572-951-2
15	109.8	4.0	2542	3	US-08-941-445A-6
16	103.4	3.7	2652	4	US-09-115-704-1
17	97.8	3.5	2671	6	5168051-9
18	96.6	3.5	1641	1	US-08-300-903A-8
19	95.6	3.5	1098	3	US-09-248-335-35
20	94.8	3.4	1117	4	US-09-247-373B-33
21	94	3.4	1582	3	US-08-545-196B-10
22	94	3.4	1582	3	US-08-545-196B-12
23	94	3.4	2852	3	US-09-027-137-2
24	92.4	3.3	1733	3	US-09-073-569-1
25	92	3.3	1464	1	US-07-735-065-1
26	92	3.3	1464	1	US-08-469-202-11
27	92	3.3	1464	2	US-08-484-434C-11

28	91.2	3.3	1813	4	US-09-071-224-3	Sequence 3, Appli
29	90.8	3.3	1882	4	US-09-370-253-1	Sequence 1, Appli
30	90	3.2	1198	3	US-09-248-335-27	Sequence 27, Appl
31	90	3.2	1474	4	US-08-821-994-64	Sequence 64, Appl
32	89.8	3.2	2303	3	US-08-836-567-1	Sequence 1, Appli
33	89.8	3.2	4168	3	US-08-836-567-11	Sequence 11, Appl
34	88.8	3.2	140	1	US-08-628-417-5	Sequence 5, Appli
35	88.8	3.2	240	1	US-08-628-417-6	Sequence 6, Appli
36	88.6	3.2	2447	2	US-09-014-969-14	Sequence 14, Appl
37	88.4	3.2	630	1	US-08-185-414E-1	Sequence 1, Appli
38	88.4	3.2	1817	1	US-08-473-981A-5	Sequence 5, Appli
39	88.4	3.2	1817	2	US-08-474-087-5	Sequence 5, Appli
40	88.4	3.2	2082	2	US-08-785-310A-2	Sequence 2, Appli
41	88	3.2	144	1	US-08-702-344-26	Sequence 26, Appl
42	88	3.2	3410	4	US-09-020-956-110	Sequence 110, App
43	88	3.2	3410	4	US-09-030-607-110	Sequence 110, App
44	87.6	3.2	1066	1	US-08-157-101A-4	Sequence 4, Appli
45	87.2	3.1	1602	1	US-08-530-950-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-09-196-390-1  
; Sequence 1, Application US/09196390  
; Patent No. 6307125  
; GENERAL INFORMATION:  
; APPLICANT: Block, Martina  
; APPLICANT: Lorz, Horst  
; APPLICANT: Lutticke, Stephanie  
; APPLICANT: Walter, Lennart  
; APPLICANT: Froberg, Claus  
; APPLICANT: Kossmann, Jens  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES  
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/196.390  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 196 21 588.9  
; FILING DATE: 29-MAY-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 196 36 917.7  
; FILING DATE: 11-SEP-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP97/02793  
; FILING DATE: 28-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley, Jr., James F.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: AGREVO-9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 596-9000  
; TELEFAX: (212) 596-9090  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2239 base pairs



[illegible]







Db	810	 AATTGGAGGATATATTTATGACAGAAATTGCATGTTTGTTCATTAATGGCATGCCA	869
Qy	1142	gccttgccagtcctctctgctgcacaaatagaccatcgcgtggtttacagagattccc	1201
Db	870	GTCTAGTGGCAGTCCCTTCCTGCTGCAANAATAGACCATATAGGTGTTTATAAGACTCCC	929
Qy	1202	gcagcaccttgtaaataattagcacatcagggtgtggagccctgcgaagtacatc	1261
Db	930	GCAGCATCTCTGTAATACATAATTTAGCACATCAGGGGTGTAGAGCCTGCAAGCACATATC	989
Qy	1262	ctgactcgggattgcctcctgaaatggatgagcctttagaattgggtatttccagaatcggg	1321
Db	990	CTGACCTTGGGTTGCCACCTGTAATGTAATGGAGCTCTGGAGTGGGTATTCCTCGAATGGG	1049
Qy	1322	caaggagcatcctctgcacaaaggtgagcagtttaaccttttgaaaggagcagtttgtga	1381
Db	1050	CGAGGAGCATGCCCTGTGACAGAGGTGAGCGATTAAATTTTGAAGGTGCAGTTGTGA	1109
Qy	1382	cagcagatcgattgtgacgcgcagtcagggtttattcctgggaggtcacaaactgctgaag	1441
Db	1110	CAGCAGATCGAATCGTGACTGTCAGTAAGGTTATTTCATGGGAGGTGCACAACTGCTGAAG	1169
Qy	1442	gtggacagggcctcaatgactgtcttaagctcccgaaaaagtgtattgaatggaaattgtaa	1501
Db	1170	GTGCACAGGGCCTCAATGAGCTCTTAAGCTCCAGAAAGAGTGTATTAAACGGAATTGTAA	1229
Qy	1502	atggaattgacattaatgattggaaccccccacacagacaagtgtccctcatcattatt	1561
Db	1230	ATGGAATTGACATTAATGATTTGGAACCTTGCACAGACAAATGTATCCCTGTCAATTATT	1289
Qy	1562	ctgtcgatgacctctctgdaaaggccaaatgttaaagctgaaatgcgaaggagtggtgtt	1621
Db	1290	CTGTTGATGACCTCTCTGGAAGGCCAAATGTTAAAGGTGCATTGCAGAAGGAGCTGGGTT	1349
Qy	1622	tacctgtaaggaggagatttctctgattggtcttatttgggaagactggattaccagaag	1681
Db	1350	TACCTATAAGGCCCTGATGTTCTCTGATTGGCTTTATTGGGAAGATTGGATTATCAGAAAG	1409
Qy	1682	gcattgatctcataaaatggccattccagactcatgagggaggaagtcgaatttgtca	1741
Db	1410	GCATTGTATCTATTTCAACTATTATACCCAGATCTCATCGGGAAGATGTTCAATTTGTCA	1469
Qy	1742	tgcttgatctggggatcccaatttttgaaggctggaatgagatctaccagagtcgagttaca	1801
Db	1470	TGCTTGGATCTGTGACCCAGAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCA	1529
Qy	1802	aggataaaattccctggatgggttggaatttagtgttccagtttccacagaataaactcgag	1861
Db	1530	AGGATAAATTTCTGGATGGTTGGATTTAGTGTTCAGTTTCCAGTTTCCACCGAATACTCCGC	1589
Qy	1862	gtgcgataattgttaatgccaatgagatttgaaccttgcggtcttaatacagctatatg	1921
Db	1590	GCTCGCATATATTTGTTAATGCCATCCAGATTCGAACCTTGTGGTGTCAATCAGCTATATG	1649
Qy	1922	ctatgcaatgtgacagttcctctgatttcagaaactcgggggctccgagacacagtcg	1981
Db	1650	CTATGCAGTATGCGACAGTTCTCTGTGTCCATGCAACTTGGGGGCCCTTAGAGATACCGTGG	1709
Qy	1982	agaccttcaacccttttgtgcacaaaggagaggggtacaggggtgggcttctcacccgc	2041
Db	1710	AGAACTTCAACCTTTCTCGGTGAGAATGGAGACAGGGGTACAGGGTGGGCATTCGCACCCC	1769
Qy	2042	taaccgtggacaagaattgttgggcaattcgcgaaccgcgagatgcgacattcgaaggagacaa	2101
Db	1770	TAAACACAGAAAACATGTTGTGGACATATCGCAACTTGCGAATATCTACATACAGGGAAACA	1829
Qy	2102	agccgtctgggggggctcatgaagcagggcagtcagcagaagaaccatcgtggagaccatg	2161
Db	1930	AGTCCTCTCTGGGAAGGGCTTAATGAACGGAGGCATGTCAAAGACTTTCAGTGGGACCATG	1889
Qy	2162	cc-ccgagcagtagtcgagcagatcttcgagtggggccttcgtggaccacaaacctcagtcagt	2220

```

Db 1990 CCGCTGAACAAATACGACAAATCTTCCAGTGGCCCTTATCGATCGACCCCTATGTCATGT 194
Qy 2221 agacgggagctggggaggt 2239
   ||  ||||  ||||
Db 1950 AAAAAGGACCAAAAGTGGT 1968

RESULT 3
US-08-572-951-1
: Sequence 1, Application US/08572951
: Patent NO. 5824790
: GENERAL INFORMATION:
: APPLICANT: KEELING, PETER L.
: APPLICANT: KNIGHT, MARY E.
: APPLICANT: GUAN, HANPING
: TITLE OF INVENTION: MODIFICATION OF STARCH
: TITLE OF INVENTION: SYNTHESIS IN PLANTS
: NUMBER OF SEQUENCES: 41
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CUSHMAN DARBY & CUSHMAN
: ADDRESSEE: Intellectual Property Group of
: ADDRESSEE: Pillsbury Madison & Sutro LLP
: STREET: 1100 New York Avenue, N.W.
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20005-3918
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/572,951
: FILING DATE: 15-DEC-1995
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/346,602
: FILING DATE: 29-NOV-1994
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/263,921
: FILING DATE: 21-JUN-1994
: CLASSIFICATION: 800
: ATTORNEY/AGENT INFORMATION:
: NAME: Paul N. Kokulis
: REGISTRATION NUMBER: 16,773
: REFERENCE/DOCKET NUMBER: 222957/1.02.15C
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 861-3000
: TELEFAX: (202) 822-0944
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2990 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: linear
: MOLECULE TYPE: cdna
US-08-572-951-1

```

	Query Match	48.1%;	Score 1332;	DB 1;	Length 2990;
	Best Local Similarity	81.1%;	Pred. NO. 1.8e-277;		
	Matches 1617;	Conservative	0;	Mismatches 340;	Indels 21; Gaps
Qy	237	cgacactctcgcgcctctccctctgccgcgcgcacacccatgcgcgcgacggcgct	296		
Db	537	CGCCATTCTTGGGGCGATACAGGGCCCCCCCCCGCGCTCTGCTGCTCCCTCT	596		
Qy	297	cggcgccgggtgcctgcgcgccccacgctccgcctgcgcgcgacccgcgcgcgcgcg	356		
Db	597	CGCGAATGGCGACGCCCTTGGCGCTGGCGCGCGCGCTGCTCTGCTGCGCGGGCGCTG	656		











QY 2088 attcaggagacacaccctcctggagggtcctcatgaagcagcagcatgacaaagacca 2147  
Db 1672 ATACAGGAAACAAAGTCTCTCTGGAGGCTAATGAAGCGAGGATGTCARAAGACTT 1731  
QY 2148 tacgtgggaccatgcccc 2165  
Db 1732 CACGTGGACCATGCCGC 1749

RESULT 5  
US-08-941-445A-20  
: Sequence 20, Application US/08941445A  
: Patent No. 6107060  
: GENERAL INFORMATION:  
: APPLICANT: Keeling, Peter  
: APPLICANT: Guan, Hanning  
: TITLE OF INVENTION: Starch Encapsulation  
: NUMBER OF SEQUENCES: 37  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
: STREET: 5370 Manhattan Circle  
: CITY: Boulder  
: STATE: CO  
: COUNTRY: US  
: ZIP: 80303  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent In Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/941,445A  
: FILING DATE: 30-SEP-1997  
: CLASSIFICATION: 800  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 60/026,855  
: FILING DATE: 30-SEP-1996  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Winner, Ellen P  
: REGISTRATION NUMBER: 28,547  
: REFERENCE/DOCKET NUMBER: 89-97  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (303) 499-8080  
: TELEFAX: (303) 499-8089  
: INFORMATION FOR SEQ ID NO: 20:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1620 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: double  
: TOPOLOGY: not relevant  
: MOLECULE TYPE: cDNA to mRNA  
: HYPOTHETICAL: NO  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: 1..1620  
: US-08-941-445A-20

Query Match 45.1%; Score 1248.4; DB 3; Length 1620;  
Best Local Similarity 86.2%; Pred. No. 1.3e-259;  
Matches 1392; Conservative 0; Mismatches 221; Indels 1; Gaps 1;

QY 553 gacgcggcgtgggggaactcgcgccgacctcctgctcgaagggtgctgaggtacc 612  
Db 4 GTCGCGGAGCTGAGCAGGAGGAGCCTCGTCTCGAACCTGGAAGGATGCTGAAGGTCC 63  
QY 613 atcgacagcataattgtggctgcaagtgcaggatcttgagatcattggatgcgaatgag 672  
Db 64 ATCGATAACACAGTAGTGTGGCAAGTGAGCAAGATTCGAGATTTGGTGGAAAGGAG 123  
QY 673 caactcaagtaaatcaactagcatcgtgtttgtgactggtgaagctgctccttat 732  
Db 124 CAAGCTCGAGCTAAAGTAACACAAAGCAATTGTCTTTGTAACCGCGAAGCTTCTCTTAT 183

QY 733 gcaaaagtcaggggggttggaagatgttggattggttaccaattgctcttctgctcgt 792  
Db 184 GCAAAAGTCTGGGGGTCCTAGGAGATGTTGGTGTTCATGCCAGTTCCTCTTCTGCTCT 243  
QY 793 ggtcaccgagtgatggttgtaatagcacaagatacttaaatgggtcctctgataaaaaactat 852  
Db 244 GGTCAACCGTGTGATGTTGTAATGCCAGATATTAAATGGTACCTCCGATAGAATAT 303  
QY 853 gcaagggaattatacacctgcgaagcacatgaagattccatgcttttgggggatacaatgaa 912  
Db 304 GCAAAATGCATTTTACACAGAAAAACACATTCGGATTCCTATGCTTTGGCGGTGAACATGAA 363  
QY 913 gtgacctttttcatgagtatagagacaacgctcgattgggtgtttgtcgtcatccgtca 972  
Db 364 GTTACCTTCTCCATGAGTATAGAGATTCAGTTGACTGGGTGTTGTTGATCATCCCTCA 423  
QY 973 tatcacagaccaggaagtttatatggagataaattttgggtctttgggtgataaatcagttc 1032  
Db 424 TATCACAGACCTGGAAATTTATATGGAGATAAGTTTGGTGTCTTTGGTGATATATCAGTTC 483  
QY 1033 agatacacactcctttgtctatgctgcagcgagggcccaactaatccttgaattggagaa 1092  
Db 484 AGATACACACTCTCTTGTCTATGCTGCATGTGAGGCTCTTTGATCCTTGAATTTGGAGGA 543  
QY 1093 tatattatggacagaattgcatgttgttgaaacgattggcatgccagccttgtgcca 1152  
Db 544 TATATTTATGGACAGAAATTCATGTTGTTGTCAATGATGGCATGCCAGTCTAGTGCCA 603  
QY 1153 gtcccttctgtcgaaaaatatagaccatacggltgttacagagattcccccgagcacctt 1212  
Db 604 GTCTTCTTCTGTCAAAAATATAGCCATATGTTGTTTATAAAGACTCCCGCAGCATCTT 663  
QY 1213 gttatacataatttagcacatcagggtgtggagcctgcagatcacatatcctatctggaa 1272  
Db 664 GTAATACATAAATTTAGCACATCAGGCTGTAGGCCCTGCAAGCACATATCCTGACCTTGG 723  
QY 1273 ttgcctcctgaaatgggtatggagcctttagaatgggtatttccagaatggcgagagggcat 1332  
Db 724 TTGCCACCTGAATGGTATGAGGCTCTGGAGTGGGTATTCCTGAATGGCGAGGAGCAT 783  
QY 1333 gcccttgacaagggtgagcagcttaaccttttgaaggaggcagctgtgacagacagatcgg 1392  
Db 784 GCCCTTGACAAGGGGTGAGGCAGTAAATTTTGAAGAGTGCAGTGTGACAGCAGATCGA 843  
QY 1393 attgtgacctcagtcaggttattcatggaggtcacaaactgctgaaggtgacagggc 1452  
Db 844 ATCGTACTGTCTAGTAAAGGTTATTCTGGGAGGTCACAACTGCTGAAGGTGACAGGGC 903  
QY 1453 ctcaatgagctcttaagctcccgaaaaagtgtattgaaatggaattgtaaatggaattgac 1512  
Db 904 CTCAATGAGCTCTTAAGCTCCAGAAAGAGTGTATTAAACGGAATTTGTAATGGAATTGAC 963  
QY 1513 attaatgattggaacccccaccacagacagagtgctccctcatcattattctgctgatgac 1572  
Db 964 ATTAATGATTGGAACCCCTGCCACAGACAAATGATCCCCCTGTCTATTCTGTGTGATGAC 1023  
QY 1573 ctctctggaaggcccaaatgaaagctgaaatgcagaagagttgggtttcacctgtaagg 1632  
Db 1024 CTCTCGAAAAGGCCAAATGAAAGGTGCATTCAGAAGGAGCTGGTGTTCCTATTAAGG 1083  
QY 1633 gaggatgttctctgattggcttatttgaagactggattaccagaaaagcattgatctc 1692  
Db 1084 CCTGATGTCTCTGCTGCTTATTGGAAGGTGGATTTATCAGAAAAGCATTGATCTC 1143  
QY 1693 attaaatggccattccagagctcatgaggaggagagctgcaaatgtgtcatgctgtgatct 1752  
Db 1144 ATTCAACTTATCATACAGATCTCATCGCGGGAAGATGTTCAATTTGTCTGCTGGATCT 1203  
QY 1753 gggatcccaatttttgaagctgagatctaccagctcaggttaccagagataaattc 1812  
Db 1204 GGTGACCCGAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCAAGGATAAATTT 1263



```
Qy 1813 cgtgatgggttgatttagtgcagttccagttccacagaataactgcaggttgatata 1872
Db 1264 CCGTGGTGGTGGATTAGTGTTCAGTTTCCACCGGAATAACTGCCGGCTGGGATATA 1323
Qy 1873 ttgttaatgccatgagattgaacctgcggtcttaatacagctatgctatcaatat 1932
Db 1324 TTGTTAATGCCATCGAATTGCAACCTTGTGCTCAATCAGCTATATGCTATGCAATP 1383
Qy 1933 ggtacagttctcttagtctcaggaactggggggtcccgagacacagtcgagacacctcaac 1992
Db 1384 GGCACAGTTCCTGTTGTCATGCAACTGGGGGCTTAGAGATACCGTGGAGAACTTCAAC 1443
Qy 1993 ccttttggtgcaaaaggagagaggggtacaggggtgaggggttctcaccgcgtacacgtggac 2052
Db 1444 CTTTCGGTGGAGATGGAGACAGGGGTACAGGGTGGGCAATTCGACCCCTTAACACAGAA 1503
Qy 2053 aagatg-ttgtgggctcattgcgaacgcgagatgcacattcagggagcacagccgtcctg 2111
Db 1504 AACATGTTTGGACATTCGCAACTGCATATCTACATACAGGGAACACAAAGTCTCCTG 1563
Qy 2112 ggaagggtcctatgaagcgaggtacgacgaagacacatacgtgggaccatgcccc 2165
Db 1564 GGAAGGGCTAATGAAGGAGGAGCATGTCAAAAGACTTCACGTGGGACCATGCGCG 1617

RESULT
US-08-836-567-9
; Sequence 9, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2360 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
```

```
; ORIGINAL SOURCE:
; ORGANISM: Solanum tuberosum
; STRAIN: cv. Desire
; TISSUE TYPE: leaf tissue
; IMMEDIATE SOURCE:
; LIBRARY: cdna-library in Lambda ZAPII
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 68..1990
; US-08-836-567-9
```

```
Query Match 27.2%; Score 752.8; DB 3; Length 2360;
Best Local Similarity 68.3%; Pred. No. 7e-153;
Matches 1088; Conservative 0; Mismatches 497; Indels 7; Gaps 3;
```

```
Qy 633 tgaagtgaacagcaggtatctgagatcatggtatgcaaatgagcaacctcaagctaaagtac 692
Db 397 TGATAGAGATGACATTGATGAGGATAGTGGAGAGATGGAGAAACCCCAATCAAAATTAAC 456
Qy 693 acgtagcatcgtgtttgtgactgtgaagctgctcctctatgcaaaagtcagggggtggg 752
Db 457 TTTCATATATCATTTTGTACTGCTGAAGCAGCTCCATATTTCTAAGACTGGTGAATTAGG 516
Qy 753 agatgtttgtgttcgtaccgaattgctctgtctgtcgtcgtggtccacgagtgatggttgt 812
Db 517 AGATGTTTGTGGTCTTTTGCATAGGCACATAGCTGCTCGGGGTCAVCGTGAATGTCGT 576
Qy 813 aatgcaagatactaaat---gggtcctctgataaaaactatgcaaaagcatatatacac 869
Db 577 TTCACCTAGGTATTTGAATGGAGGTCTTCAGATGAAAGATAGCCCAATGCTTTGACCT 636
Qy 870 tgggaagcacatlaagattccatgctttggggatcacatgaaagtgcacctttttcatga 929
Db 637 TGATGTCGGGGCCACATGTCATTGCTTGGTGTGATGCACAGAGAGTACCCCTTCTACCATGA 696
Qy 930 gtatagagacaacgctcgatgggtgtttgtcgatcatccgtcatatcacagaccaggaag 989
Db 697 ATACAGGGCAGGTGTTGATGGGTATTTGTGGACCACCTCTTCTACTGCAGACCTGGAAC 756
Qy 990 ttatatggagataatttgggtgcttttggatataatcagttccagatcacacaccccttg 1049
Db 757 GCCATATGGTATTTATGTCGATTTGGTGAATACATGATTCGGCTTCATCTTGTGCTTTC 816
Qy 1050 ctatgctgcacgagggccccactaaaccttgaattggaggagataattttatggacagaa 1109
Db 817 TCACGCGAGCATGTGAAGCGCCATTGGTCTTCCACCTGGGAGGGGTTCATTTATGAGAGAA 876
Qy 1110 ttgcatgtttgtgaaacgattggcatgccagccttgtgccagtcctcttctgtgcaaa 1169
Db 877 GTGCTTGTTCGCTAAATGATTGGCATGCTGCCCTGGTTCCTTTTACTTTTAGCGGCCAA 936
Qy 1170 atatagaccatacgtgtttacagagattccgcgacagcccttggtatacaataattagc 1229
Db 937 GTATCGTCTTATGGTGTTTACAGAGATGCTCGTAGATTATTCGCCAATACACAAATTCG 996
Qy 1230 acatcaggggtgtgagcctgcgaagtacatctatctgattcgtgggattgcctcctgaaatga 1289
Db 997 ACATCAGGGAGTGGAGCCTGCAGTAACCTACAATAATTTGGGTTTGGCTCCCAATGGTA 1056
Qy 1290 tggagctttagaatgggtatttccagaaagggcagagagcatgcccttgacaaaggtga 1349
Db 1057 TGGAGCAGTGAATGATATTTCCACATGGGCAAGGGCGCATGGCGCTTGACACTGGTGA 1116
Qy 1350 ggcagttaacttttgaaggagcagttgtgcagcagatcggattgtgaccgtcagta 1409
Db 1117 AACAGTGAACGTTTTTGAAGGGGCAATAGCAGTTGCTGATCGATACGATCTACAGTTAGCCA 1176
Qy 1410 ggggtattcatgggaggttcacacactgctgaaggtggacagggcctcaatgagctcttaag 1469
Db 1177 GGGATACTCATGGGAATAAACAACTCCTGAAGGGGATATGGCGGTACATGAGCTGTTGAG 1236
Qy 1470 ctcccgaaaaagtgattgaaatggaattgaaatggaattgacattatgattggaaccc 1529
```



Db 1237 CAGTACAGAGTCTGTCTTAAATGGAATTAATGAATAGATGTTAAATGATTGGAACCC 1296  
QY 1530 caccacagagaagtgtctccctcaccattctctgcatgacctctctggaaagccaa 1589  
Db 1297 GTGACAGATGACATATTTCTTGGCATTTACCTCATTAATGACCTCTCCGAAAGGTTCA 1356  
QY 1590 atgtaagctgaattgcagaagagtggtgtttacctgtlaagggagagtgcttccctgat 1649  
Db 1357 GTGCAAGACTGATCTGCAAAAGGAAGTGGCCCTTCAATTCGACCTGATTGCTCTGAT 1416  
QY 1650 tggctttattgagaagactgattaccagaagaagcattgtatctcattaaatggccattcc 1709  
Db 1417 TGATTTATTGGAAGGCTGGACTACAGAAAGCTGTGACATAATCTCTCAGCAATTC 1476  
QY 1710 agagctcatgaggaggagtgcaattgtcattgtctgtatctggtatggatcccaattttga 1769  
Db 1477 AGAAGTATGACAGATGATGCTCAAGTGTATGCTTGGATCTGGTGAAACAATATGA 1536  
QY 1770 aggttgatgagatctaccagtcagtcagttacaaagataaattccgtggatgggttgatt 1829  
Db 1537 AGACTGGATGACATACAGAAATCTTTTAAAGACAAATTTCTGTGTTGGTTGGATT 1596  
QY 1830 tagttccagtttcccaagataactgcaggttgcgtatctatattgttaatgccatcgag 1889  
Db 1597 TAATGTTCCAGTTTCTCATAGGATAACAGCAGGATGCGACATACACTATTGATGCCCTCAAG 1656  
QY 1890 attgaaccttcggtcttaacatcagctatctgctatgctgaatgcaatgatacagttccctgagt 1949  
Db 1657 ATTCGACCGTGTGGCTTAAACCAATTTGATGCAATGAGATATGACCATACCTATTGT 1716  
QY 1950 tcatggaactggggcctccgagacacagtcagagaccttcaaccttttgggtgcaaaagg 2009  
Db 1717 TCATAGCAGCGGGGCTTAAGAGACACAGCTGAAGATTTTAATCCATATGCTCAAGAGG 1776  
QY 2010 ---agaggaggtacaggtgggggtctccaccgttaaccgttggacaagatgttggtggc 2066  
Db 1777 ATAGGTGAAGTACCGGGTGACATTTTCTCTCTAAGCAGTGAAGATTGCTTGTATAC 1836  
QY 2067 attgcaaccgcatgtgcacattcagggagcacagcgctctctggaggggctcatgaa 2126  
Db 1837 ACTGAAGTGGCAATCGGGACTTATACAGACATAGTCAATCTGGAGGGGATTGATGAG 1896  
QY 2127 gcgagcatgacgaagaccatacagtgggaccatgc-cgcgagcagtcagcagcatctt 2185  
Db 1897 GAGAGGTATGGAAGGAGCTATTCTCGGAAATGCAGCAATTCATATGAACAAGTTT 1956  
QY 2186 cgaagtgcccttcgaggaccaacctacgtca 2217  
Db 1957 CACCTGGGCTTTATAGATCCTCCATATGTCA 1988

RESULT 7  
US-08-836-567-3  
; Sequence 3, Application US/08836567  
; Patent No. 6130367  
; GENERAL INFORMATION:  
; APPLICANT: Kossmann, Jens  
; APPLICANT: Springer, Franziska  
; APPLICANT: Abel, Gernot  
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES  
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/836,567  
; FILING DATE: 24-JUL-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP95/04415  
; FILING DATE: 09-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 44 41 408.0  
; FILING DATE: 10-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley Jr., James F.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: Agrevo-4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-596-9000  
; TELEFAX: 212-596-9090  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1758 base pairs  
; TYPE: nucleotide  
; STRANDEDNESS: unknown  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Solanum tuberosum  
; STRAIN: cv. Berolina  
; TISSUE TYPE: tuber tissue  
; IMMEDIATE SOURCE:  
; LIBRARY: cDNA-library in pBluescriptSKII+  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1377  
; OTHER INFORMATION: /function= "Polymerization of  
; OTHER INFORMATION: starch"  
; OTHER INFORMATION: /product= "Starch synthase"  
US-08-836-567-3

Query Match 23.4%; Score 648.4; DB 3; Length 1758;  
Best Local Similarity 68.6%; Pred. No. 1.8e-130;  
Matches 941; Conservative 0; Mismatches 421; Indels 10; Gaps 3;  
QY 856 aaggcattatcacactgcgaagcacatttaagattcccatgctttgggggagcacatgaagtg 915  
Db 10 AATGCTGTGACCTTGATGCGGGCCACTGTCCATTGCTTGGTGATGCACAGGAAGTA 69  
QY 916 accctttttcatgagtagagacaacgtcgattgggtgtttgtcgcataatccgtcatat 975  
Db 70 GCCTTCTACCATGAATACAGGCGCAGGTGTGATTGGGTATTGTGGACCACTCTCTTAC 129  
QY 976 cacagaccagaagtttatagagataatttgggtgcttttgggtgataaatacagttcaga 1035  
Db 130 CCAGACCTGGAACCCCATATGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 189  
QY 1036 tacacactccttgcattgctgcagcgagcccccactaaactccttgaattggagagatat 1095  
Db 190 TTCACCTTTGCTTCTCACGACGACATGTGAGCGCCATTGGTCTTCCACTGGAGGGTTC 249  
QY 1096 atttatgacagaattgcatgtttgtgtgaacgattggcagcgttgccagcgttgccagtc 1155  
Db 250 ACTTATGGAGAGAGTGTGTTTCTCGTAATGATTGCAACGCTGCCTTGGTCTCTTTA 309  
QY 1156 cttcttgcgtcaaaatagaccatacagtggttttacagagattcccgagcagaccctgtt 1215  
Db 310 CTTTATGGCGCCAGTAGTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 369  
QY 1216 atacataatttagcacatcagggtgtggagcctgcaagtcacatcatcctgatctgggattg 1275



Db	370	ATACACAACTTCACATCAGGAGTGGAGCCTGCAGTAACCTACAATAATTTGGGTTG	429
Qy	1276	cctcctgaatggtatggagctttagaattgggtatttcagaatagggaagagcagtcgc	1335
Db	430	CTCCACAATGGTATGGACAGTTGAATGGATATTTCCACATAGGCAAGGCGCATCGC	489
Qy	1336	cttgacaagggtgaggcagtttaacatttttgaagaagagcagttgtgacagcagatcggatt	1395
Db	490	CTTGACACTGGTGAACAGTAGAACGTTTGAAGGGGCAATAGACAGTTGCTGATCGGATA	549
Qy	1396	gtgacctgaactcagggttattcatgggaggttcacaactcgtgaaggtggacagggcctc	1455
Db	550	CTGACAGTTAGCCAGGGATACTCATGGGAAATAACAACCTCCTGAAGGGGATATGGGCTA	609
Qy	1456	aatgagctcttaagctcccgaaataaagtgtattgaatggaattgtaaatggaattgacatt	1515
Db	610	CATGAGCTGTTGACGAGTAGACAGTCGTCTCTTAATGGAATTAATAATGGAATAGATGTT	669
Qy	1516	aatgattggaaaccccccagacagaagtgtctccctcatcattctgtctcgaagcctc	1575
Db	670	AATGATGTTGGAACCCGTCGACAGATGACATATCGCTTCGCATTACTCCATCAATGACCTC	729
Qy	1576	t-----ctgaaaggccaaatgtaagctgaattgcgaaggaggttgggtttacctgta	1629
Db	730	TCCCCCCTCGGAAGGTTGAGTGCAAGATGTATGCAAAAAGGAACCTGGCGTTTCCAATT	789
Qy	1630	agggagagattcctctgattggcttattgaaagactggaattaccagaagaagcattgat	1689
Db	790	CGACCCGATTTGCCATGATTTGATGGATTTATGGAGGCTTGGACTACCAAGAAAGTTGTGAC	849
Qy	1690	ctcattaaatggtccattccagagctcatgagggaggagctgcgaatttgtcatgcttggga	1749
Db	850	ATAATCTGTGCAGCAATTCAGAACTTATGCAGAATGATGTCCAAGTTGTAACTCTTGA	909
Qy	1750	tctgggagtcacaattttgaagcgtgagatgagatctaccgagtcgagttacaagagataaa	1809
Db	910	TCTGGTGAGAAACAATATGAAGACTGGATGAGACATACAGAAAATCTTTTTAAGACAAA	969
Qy	1810	tccgctggaatgggttggattagttgctcagtttcccacagaaataactgcaggttgcgat	1869
Db	970	TTTCGTCTTGGGTGGATTTAATGTTCCAGTTTCTCATAGGATAACAGCAGGATGCGAC	1029
Qy	1870	atatgttaatgcaatcgagattgaaccttcggtcttaatacagctataatgctatgcaaa	1929
Db	1030	ATACTATTGATGCCCTCAAGATTCGAACCGTGTGGCTTAAACCAATGTATGCAATGAGA	1089
Qy	1930	tatgttacagtctcctgtagtctaatgaaactggggcctccagacacagtcgagaccttc	1989
Db	1090	TATGGCACCATACCTATTGTTTCATAGCAGCGGGGGCTTAAGAGACACAGTCAAGGATTTT	1149
Qy	1990	aaccttttgggtcaaaagg---agagagggtacagaggtggggcttctccacgcctaacc	2046
Db	1150	AATCCATATGCTCAAGAAGGAAAGGTGAAGTACCGGGTGGACATTTTCTCCTCAACG	1209
Qy	2047	gtggacaagatgttggggcatttgcgaaccgagatgcgaattcaggagagcacaagccg	2106
Db	1210	AGTGAAAGTTGTTTGATACACTTGAAGCTGGCGGATCAGGACCTTATACAGACATTAAGTCA	1269
Qy	2107	tccctgggaggggtctcatgaagcagggcatgacgaagaagcacaacgtgggacccatgc-ccc	2165
Db	1270	TCTTGGCAGGGATTGATGAAGAGAGAGGTATGGGAAGGGAATAATCTCTGGGAAATGCAGCC	1329
Qy	2166	gagcagttacagacagatcttctcagttggggccttcgtggaccacaacctacgtca	2217
Db	1330	ATTCAATATGACAGAGTTTTCACCTGGGCTTTATAGATCTCTCCATATGTCA	1381

## RESULT 8

US-08-572-951-3

03 00 374 331 ; Sequence 3, Application US/08572951

; Patent No. 5824790

; GENERAL INFORMATION:

APPLICANT: KEELING, PETER L.  
 APPLICANT: KNIGHT, MARY E.  
 APPLICANT: GUAN, HANPING  
 TITLE OF INVENTION: MODIFICATION OF STARCH  
 TITLE OF INVENTION: SYNTHESIS IN PLANTS  
 NUMBER OF SEQUENCES: 41  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: CUSHMAN DABY & CUSHMAN  
 ADDRESSEE: Intellectual Property Group of  
 ADDRESSEE: Pillsbury Madison & Sutro LLP  
 STREET: 1100 New York Avenue, N.W.  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20005-3918  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/572,951  
 FILING DATE: 15-DEC-1995  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/346,602  
 FILING DATE: 29-NOV-1994  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/253,921  
 FILING DATE: 21-JUN-1994  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Paul N. Kokulis  
 REGISTRATION NUMBER: 16,773  
 REFERENCE/DOCKET NUMBER: 22957/1.02.15C  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 861-3000  
 TELEFAX: (202) 822-0944  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2380 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-572-951-3

## RESULT 8



Qy	997	ggagataaatcttgggtgcttttggtgataatcagtttcagatatcacacacctccctttgctatgcct	1056
Dy	998	: :	1057
Dy	999	AAATATTAYTGGGNGGARMGNNTNGAYATHYTNAAARMGNATGATHTYTNNTYTGYAARGC	1014
Qy	1007	gcattcgaggccccactaatcctgaattggagagatatattatggaca--gaattgc	1113
Dy	1008	: :	1114
Dy	1015	GCNGTNGARNTNCCNTGGTAYGCNCCTGYCGNGNACNGTNTAYGNGAGYGNAAYYTN	1074
Qy	1114	atgttttgtgaacagtattgcagccttgtgccagtccttgtgccagtccttgtctgctgcaaaatat	1173
Dy	1075	GTNTTATTHGCNAAAYGAYTGCCAYACNCNNTYNTCNCNTGTAYTNAARGCNTAYTAY	1134
Qy	1174	agaccatacgggtgtttcacagagattcccgcagcaccttgttatatacataaatttagcacat	1233
Dy	1135	MGNAYAAAYGGNYTNATGCARTAYGCNMGNNSGNTYNTGTNATHCAAYAYATHGCNCAY	1194
Qy	1234	cagggtgtgagccctgcaagtacatatcctgatctggattgacctcctgtaagtggatcgga	1293
Dy	1195	CARGNMNGNCCNG-TNGAYGATTTGTNAAYTYTGAYTNNCCNGARCAYATAYATHGA	1253
Qy	1294	gccttagaattgggtatttccagaatgggcaaggagcgatgcccttgacaagggtgagcca	1353
Dy	1254	YCAYTTYARYTNAYGAYAAATHHG-----NGNGAYCAY	1290
Qy	1354	gttaaccttttgaaggagcagttgtgacagcagatcgattgtgacgtcagtcagcggt	1413
Dy	1291	WSNAAVGTNTTYGCNCNGNNTNAAACNGCNGAYMGNTGTNACNTGMSNAAVGGN	1350
Qy	1414	tattcatggagggtcacaaactgctaagggtggacagggcctcaatgagctctttaagctcc	1473
Dy	1351	TAYATGTGGGARVTNAAACNMSNARGNGGNTGGGNYTNCAYGAYATHATHAAYCAR	1410
Qy	1474	cgaaaagtgattgaattggaattgaattgaattgacattaatgattggaaccccaacc	1533
Dy	1411	AAYCATYGGAAARYTNCARGNATHGTNAAVGGNATHGAYATGWSNGARTGGAAYCCNGCN	1470
Qy	1534	acagacaagtgctcctcctcatattctgtcgatgacctct-----ct	1578
Dy	1471	GTNGAYGTNCAYTNCAYWSNGAYGAYTAYACNAAAYTAYACNTTYGARACNYTNGAYACN	1530
Qy	1579	ggaaaggccaaatgtaaagctgaattgcagagaggagttgggtttacctgtaaagggaggt	1638
Dy	1531	GGNAAARMGNCARTGYAARGCNCNTNCARMGNCARYTNGGNYTNCARTGMNGNGAYGAY	1590
Qy	1639	gtctctctgattgcttattggaactggaattaccagaagcattgactcattcaataa	1698
Dy	1591	GTNCCNYTNATHGNTTYATHGGNMGNTNGAYCAYCARAARAAGNGTNGAYATHATHGCN	1650
Qy	1699	atggccattccagagctcatgaggaggaagcgtgcaattgtcgtgctggtgctggggat	1758
Dy	1651	GAYGCNATHCAYTGGATHGCNGCNGCARGAYGTNCARYTNGTNTATGTYTNGGNACNGMGN	1710
Qy	1759	ccaatttttgaagcttgatgagatctaccagatcgagttcacagggataaaattccgtgga	1818
Dy	1711	GCNGAYTNGARGAYATGYTNMGNTTYGARWSNGARCAYWSNGAYAAARGTMNGMNGCN	1770
Qy	1819	tgggttgatttagtttccagttlccacagataaactcgagttcgatatattgtta	1878
Dy	1771	TGGGTNGNTTYWSNNTNCCNTNGCNCAYMGNATHACNGCNGNGNGNGAYATHATHYTN	1830
Qy	1879	atgccatcgagatttgaaccttgcggtcttaatcagctatctgctatgcaataatggtaca	1938
Dy	1831	ATGCCNWSNMNTTYCARCCTGYGNYTNAAYCARYTNTAYGCNATGGCNTAYGGNACN	1890
Qy	1939	gttccctgatttcaatgaaactggggcctccgagacacagtcgagaccttcaaacctttt	1998
Dy	1891	GTNCCNGTNGTNCAYCNGTNGSGNNTYTMNGNGAYACNGTNGCNCNTTYTGAYCCNTTY	1950
Qy	1999	ggfcaaaaaggagagggttacagggtggggcggttctcaccgcgtaaacgctggacaagatg	2058
Dy	1951	AAY-----CAYACNGGNTNGNTGGACNTTYTGAYMNGCNGARGCNAAAATMGNATG	2001
Qy	2059	tgtgggcatctgcgaaccgcgatgtcgacattcagggagacaaagcgtcctcgagagggg	2118

```

Db 2002 ATHGAYCNCNTNNSNCAYTGYTTCNACNACNTAYTNGNAAYTAYAAARGARWNTGGMGNCN 2061
Qy 2119 ctcatgaagcagcgatgcgaagaacaccatcagctgggaccatgcctccc 2165
Db 2062 TGYMNGCNCNGNGNATGGCNGARGAYTTNNWSNTGGGAYCAYCGNGC 2108

RESULT 9
US-08-836-567-7
; Sequence 7, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2793 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Solanum tuberosum
; STRAIN: cv Dsire
; TISSUE TYPE: leaf tissue
; IMMEDIATE SOURCE:
; LIBRARY: cdna-library in Lambda ZAPII
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 242..2542
; US-08-836-567-7

```

Query Match  
10.48; Score 287; DB 3; Length 2793;

Best Local Similarity 53.98; Pred. No. 1.3e-52;

Matches 801; Conservative 0; Mismatches 605; Indels 81; Gaps 7;

Qy 697 agcatcgtgttttgtgactggtgaagctgctccttatgcaaaagtcagggggggttgggagat 756



Db	1070	AACATTATTTTGGTGGTTTCAGAAATCCGCTCCATCGTCTAAACAGAGTGGCGCTTGGAGAT	1129
Qy	757	gtttggtgcttaccaaatgctcttctgctctgctggtccaccgagtagtggttgtaagt	816
Db	1130	GTTGCTGGAGCAATTACCCAAGCTTTGGCTGACGCTGGCCACAGATTATGTTGTGTGCCA	1189
Qy	817	ccaagatcacttaaatgggtccctctgataaaaaactatgcaaaaggcaattatcacactgcgaag	876
Db	1190	CTTCGGTT-----ATGCAAACTATTCCTGAACCTCAAGATTTCTGGTGTA	1231
Qy	877	cacattaagattcccatgctttgggggacacatgaagtgaacctttttcatgagatataga	936
Db	1232	AGAAAAATTTATAAGTTGATGGTCAGGATGTGGAAGTGACCTTACTTCCAAGCTTTTATT	1291
Qy	937	gacaacgtcgaattgggtggtttgtcga---tcactcgtcatatcacagaccaggaagtta	993
Db	1292	GATGGTGTGGAATTTGTTTTCATTTGACAGTCATATGTTTACACACATTTGGGAACAACATT	1351
Qy	994	tatggagataaatttgggtgcttttgggtgataatcagtttcagatcacacactccttgcctat	1053
Db	1352	TACGAGGGAACCGTGGGATATTTAAACGCATGGTTTAT-----TTTGGCAA	1402
Qy	1054	gctgcatacagagggcccaactaatcccttgaattgggagatatattatggaca---gaat	1110
Db	1403	GCAGCGATTGAGGTTCCTTGGCATGTTCATCTGGTGGGGTCTGCTATGGAGATGGAAAT	1462
Qy	1111	tgcatgtttgttgcaacgttggcatgccagccttgccagtccttcttgctgcgaaa	1170
Db	1463	TTAGTGTTCATTGCTTAATGATTGGCANATCTGCTTTATTGCCAGTATATCTGAAAGCTTAT	1522
Qy	1171	tatagaccatacgggttttacagagattcccgacgaccccttgttatcacaaatttagca	1230
Db	1523	TATCGTGACAATGGAATTATGAACATATACAAGATCTGCTGGTGATTATACATACGCT	1582
Qy	1231	catcaggggtggagcctgcgaagcacatatcctgacttgggattgacctcctgaaatggat	1290
Db	1583	CATCAGGGTCGTGGTCCCTTTGGAGGATTTTCATATGTAGATCTTCCACACACTATPATG	1642
Qy	1291	ggagctttagaatgggtatttccagaatgggcaaggagcattgcccttgacaagggtgag	1350
Db	1643	GACCTTTCAAGTTGTATGACCAGTAGGA-----GGTCAG	1678
Qy	1351	gcagttaaacttttgaaggagcagttgtagacgacagatcggaattgacctgcagtcag	1410
Db	1679	CATTTCAACATTTTTCGGCTGGCTTAAGACAGACAGATCGTGTAGTTACAGTTAGTCAT	1738
Qy	1411	ggttattcatgggaggtcacaactcgtcaaggtgacagggcctccaatgagctcttaagc	1470
Db	1739	GGATATTCATGGGACTAAAGACTTCCCAAGGTGTCGGGATTGCATCAGATAATTTAAT	1798
Qy	1471	tccgaaaaagtgtattggaatggaattgaaattgacattatgatttggaaacccc	1530
Db	1799	GAGAACGATTGGAAATTTACAGGTATTGTGAATGGGATTGATACAAAGAGTGGAAACCT	1858
Qy	1531	accacagacaagtgtctccctcatcatattctgtcgatgacctct-----	1576
Db	1859	GAGTTGGACGCTTCATTACAGTCACATGGTTACATGAACTACTCCTTGGACACGCTACAG	1918
Qy	1577	-ctggaagagcccaaatgaagctgaaattgcagaaggagttgggtttacactgtaaggag	1635
Db	1919	ACTGGCAACGCTCAATGTAAGCTCATTTGCAGAGGAACCTTGCTTTTACCAGTTCGTGAT	1978
Qy	1636	gattgtcctctgattggctttattggaagactgattaccagaaggcattgatctcatt	1695
Db	1979	GATGTCCCACTGATCGGTTTTCATGGGAGGCTTGACCCACAAAAAGGCTGTGATCTGATT	2038
Qy	1696	aaatggccattccagagctcatgagggagcagctgcaatttgcactgtctgattcgtgg	1755
Db	2039	GCTGAGGCCAGTGCCTTGGATGATGGGTGAGGATGTACAACTGTGTCATGTTGGGGACGGG	2098
Qy	1756	gatcccaatttttgaaggcttggatgatagatctaccgcgtcgagttcacaggataaattccgt	1815

Db 2099 AGGCGTGACCTTGAAACAGATGCTAAGGCAATTTGAGTGTCAACAATGATAAAATTAGA 2158

Qy 1816 ggaatgggttgatttagttgccagtttccacagaataaactgcaggttgcgatatattg 1875

Db 2159 GGATGGGTTGGTTCTCTGTGAAGACTTCTCATCGTATAACTGCTGCTGCAGACATTTCTG 2218

Qy 1876 ttaatgcatacgagatttgaaccttgcggttcttaatacagctatatacgatgcgaatatgggt 1935

Db 2219 CTCATGCGCTTCTAGATTGAGCCTTGGGGACTGAACACAGCTTTATGCAATGAATATCGG 2278

Qy 1936 acagtctcgttagttcatggaactggggcctccgagacacacagtcgagaccttcaaccct 1995

Db 2279 ACTATTCTGTGTTCATGCGAGTAGGAGGACTCAGACATACTGTGACAGCCCTTTGATCCT 2338

Qy 1996 ttggtgcaaaagagagagggttacaggggtggcgcttccaccgctaaacgctggacaag 2055

Db 2339 TTTAATGAGTCAGACTG-----GGGTGGACCTTCAGTAGGGCTGAAGCTAGCCAG 2389

Qy 2056 atgtgtgggcattgcgaacccgcatgcacatctcaggaagcacagaagccgtccctggag 2115

Db 2390 CTGATCCAGCGATTAGAAATTTGCTTACTGACTTATCGTGAGTACAAAGAGTTGGGAG 2449

Qy 2116 gggctcatgaagcaggcagtcacgaagacacatacgtgggacctgc 2162

Db 2450 GGGATTACAGACACGTTGTATGCACACAAGACTTAAGTTGGGATAATGC 2496

RESULT 10

US-08-836-567-5

Sequence 5, Application US/08836567

Patent No. 6130367

GENERAL INFORMATION:

APPLICANT: Kossmann, Jens

APPLICANT: Springer, Franziska

APPLICANT: Abel, Gernot

TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES

TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC

TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: FISH & NEAVE

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/836,567

FILING DATE: 24-JUL-1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP95/04415

FILING DATE: 09-NOV-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P 44 41 408.0

FILING DATE: 10-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Haley Jr., James F.

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: Agrevo-4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-596-9000

TELEFAX: 212-596-9090

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1926 base pairs

TYPE: nucleotide

STRANDEDNESS: unknown



%	TOPOLOGY:	linear
%	MOLECULE TYPE:	CDNA to mRNA
%	HYPOTHETICAL:	NO
%	ANTI-SENSE:	NO
%	ORIGINAL SOURCE:	:
%	ORGANISM:	Solanum tuberosum
%	STRAIN:	cv. Berolina
%	TISSUE TYPE:	tuber tissue
%	IMMEDIATE SOURCE:	:
%	LIBRARY:	CDNA-library in pBluescriptSK+
%	FEATURE:	:
%	NAME/KEY:	CDS
%	LOCATION:	2..1675
%	OTHER INFORMATION:	/function= "Polymerization of starch"
%	OTHER INFORMATION:	/product= "Starch synthase"
%	US-08-836-567-5	:

  

Query Match	10.1%	Score 279;	DB 3;	Length 1926;
Best Local Similarity	53.5%;	Pred.	No. 5.8e-51;	
Matches	796;	Conservative	0;	Mismatches 610; Indels 81; Gaps

  

QY	697	agcattcgctgtttgactggtagtcgcctcttatacaaaagtccagggggttgaggagat	756
Dn	203	AACATTATTTTGTGGCTTCAGAAATGCGCTCCATGGTCTAAAACAGSGTGCGCTTGGAGAT	262
QY	757	gtttgsgtttcgttaccaattcctcttgtctcgtcgtgccaccgaagtcatggttctaagt	816
Dn	263	GTTGCTGGAGCATTACCACAAGCTTGGCTGCACCTGCCACAGAGTTATGCTTGTGGCA	322
QY	817	ccaagaatactaaaatgggttcctctgatataaaactatgcaaagggcattatatcacctgcgaag	876
Dn	323	CCTCGTT-----ATGACRACTATCTGAACCTCACAGATTCCTGGTGA	364
QY	877	cacattaagattcccacgtctttgggggataacaatgaagtgccttttttcagtatagata	936
Dn	365	AGAAAAAATTAATAAGTTGATGGTCAGGATGGGAAGTAGCTACTTCCAAGCTTTTATT	424
QY	937	gacacgcgcgattgggtgtttgtcga----tcacocgctcatacacagaccaggaagtta	993
Dn	425	GATGGTGTGGATTTTGTTTTTCATTGCACAGTCATATGTTTACACACATTTGGGAACAACATT	484
QY	994	tatggagataaatttgggtgcttttgggtgataatcagttccagtcacacctctcttgcctat	1053
Dn	485	TACGGAGGAGAACC GTGTGGATATTTAAACGCATGTTTTAT-----TTTGC AAA	535
QY	1054	gctgcacgcgagccccactaatccttgaattggaggatataatttatggaca---gaat	1110
Dn	536	GCAGCGATTGAGGTTCCCTTGGCATGTTCCATGTGTGGGGGTCTGCTATFGGAGATGGAAT	595
QY	1111	tgcattgttctgtgaacgattggcatccagcccttctgcagtcctcttctgtgcacaaa	1170
Dn	596	TTAGTCTTCAATTGCTAATGATTTGGCATAC TGCTTTATTTGCCAGTATATCTGAAAGCTTAT	655
QY	1171	tatagaccatacggfittacagagattcccgcagcacaccttgttatacataatttagca	1230
Dn	656	TATCTGTCAANTGGAAATTATGAACCTATAAAGATCTGCTCTGGTGATTCTAACATCGCT	715
QY	1231	catcaggggtgtggagccctgcagtgacatactcctgatctgggattgcctcctcgaatggat	1290
Dn	716	CATCAGGGTCTGGTCTTTTGGAGGATTTTTCATATGTAGATCTTCCCACCACTATATG	775
QY	1291	ggagctttagaatgggttattccagaataggccaaggagccattgaccaagggtyag	1350
Dn	776	GACCCYTTCAAGTTGTATGACCCAGTAGGA-----GGTGAG	811
QY	1351	gcagtttaaactttttgaaaggagcagttgtgcacagatcgattgtgacogtcagtcag	1410
Dn	812	CATTTCAACATTTTTCGGCTGGTCTAAAGACACGACATCGTGTAGTTACAGTTAGTCAT	871
QY	1411	ggttattcatggaggttcacaactgctgaaaggttgagaggccctcaatgagctctttaagc	1470



Query Match	8.4%	Score 233:	DB 3:	Length 2097;
Best Local Similarity	52.5%	Pred. No. 4.8e-41;		
Matches 796;	Conservative 0;	Mismatches 645;	Indels 76;	Gaps 10;
QY	697	agcatcggtttgtgactggtggaagctgctcccttatgcaagtcaggggggttgggagat	756	
Db	622	AACGTGCTCGTGTGGCTTCTGAATGTGCTCTTCGCAAGACAGGTGGCGCTTGGAGAT	681	
QY	757	gtttggttcgttaccaattgcttcgttcgttcgttcgttcacgcagtgatggttgtaatg	816	
Db	682	GTGCTGGGTGCTTTGGCTTAAGCTCTCGCAGGAGAGACACCGTGTATGGTCGTGNTA	741	
QY	817	ccaagataccttaaatgggttcctctotgataaaaactatgcaaggcattatacactgcgaag	876	
Db	742	CCAAGATATATGA-----GAGTATGCCGAAGCCCGGGATTTAGGTGTAAAGAGA	789	
QY	877	cacatlaagattccatgctttgggggacacatgaagtgcactttttcttgaagtataga	936	
Db	790	CGTTACAAAGTAGTGACAGGATTCA-----GAAGTTACTTATTTTCCTCTTACATT	843	
QY	937	gacacgtcgatgggttggttggatgcatactgcgtcatcacagaccaggaagtttatat	996	
Db	844	GATGGAGTTGATTTGTATTCGTAGAAGCCCTCCCT-TCGGGACCGCGCAATAATAT	902	
QY	997	ggagataatttgggtgcttttgggtataactcagttcagatacacacactcctttgctatgct	1056	
Db	903	-----TTATGGGGGAGAAAGATTGGATATTTGAAGCGCATGATTTGTTCTGCAAGGCC	957	
QY	1057	gcattcgagggcccaactaaactccttggaattgggagataattttatggaca---gaattgc	1113	
Db	958	GCATGTGAGGTTCATCGGTATGCTCCATGTGGCGGTACTGCTATGTGTATGGCAACATTA	1017	
QY	1114	atgtttggttgtaacgattggcatgccagccttggccagtcctcttcttgcgtgcaaaat	1173	
Db	1018	GTTTTCATTGCTAATGATTGGCATACCGCACTTCTGCTGCTATCTCTAAAGGCGCTATTAC	1077	
QY	1174	agaccatacgggtgtttacagagattcccgagacaccttggtatacacataaatttagcacat	1233	
Db	1078	CGGACAATGGTTTGAATGACAGTATGCTGCGCTCTGTGTTGTATACACAACTATGCTCAT	1137	
QY	1234	cagggtgtgagcctgcaagtacataactctgatctggagattgcctcctgaatggatgga	1293	
Db	1138	CAGGTCGTGCGCTGTAGACAGCTTCGTCATTAATTTTGACTTTCGCTGAACTAC-ATCGA	1196	







[illegible]

## RESULT 13

```

US-08-941-445A-8
; Sequence 8, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Haining
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2007 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2007
US-08-941-445A-8

```







```

Db 1091 AACGACTGGAAGATCAATGGCATTCGCTGAACGCATCGACACACAGGAGTGAACCCCAAG 1150
Qy 1534 acagacagtgctct-----ccctcatcattattctgtcgatgaactctct 1578
Db 1151 GTGACGTCACCTCGCGTGGAGCGGCTACACCAACTACTCCCTCGAGACACTCGACGCT 1210
Qy 1579 gaaagccaaatgaagctgaattgcagaagagtggttaccctgtgaaggagagat 1638
Db 1211 GAAAGCGGAGTGAAGCGGCGCTCGACGCGGACGCTGGAGCGCTGGAGTGGCGGACAC 1270
Qy 1639 gttctctgattggtcttattggaagactgattaccagaagacattgattcattaaa 1698
Db 1271 GTGCGGCTGCTCGCTTCATCGCGGCTTGATGATGACAGAGGCGTGGACATATCGGG 1330
Qy 1699 atggccattccagagctcatgagggagagctgcaatttgcattgcttggtgagat 1758
Db 1331 GACGCGATGCGCTGATCGCGGGGACGACCTGCAGCTGCTGCTGGCGACCGGCCA 1390
Qy 1759 ccaatttttgagctgagatctacagatctacagagctcgagttacaagataaaattcgtgga 1818
Db 1391 CTGACCTGGAACGAATGCTGCAGACTTGGAGCGGGAGCATGCCAACAAAGTGGCGGG 1450
Qy 1819 tgggttgatttagttccagtttccacagaataaactgcaggttgcgatatattgta 1878
Db 1451 TGGTGGGTTCTCGGCTCTAATGTCATCGCATCACGCGGGCGCCAGCGTCTGGTG 1510
Qy 1879 atgcacatgagattgaaccttgcggttcttaatacagctatatgtatgcaatatgtgaca 1938
Db 1511 ATGCCCCCTCCCTTCG---CGCGGGGCTGAACAGCTCTACGCGGATGGCATACGCGAC 1567
Qy 1939 gttctgtagttcagaaactggtggtggtggtggtggtggtggtggtggtggtggtggt 1998
Db 1568 GTCCCTGTGTGTCACGCGGTGGCGGCTCAGGGACACCGTGGCGGCTTCGACCCGTT- 1626
Qy 1999 ggtgcaaaagagagaggggtacaggggtggggttctcaccgctaaacgctggacaagatg 2058
Db 1627 -----CGCGACGCGCGGCTCGGCTGGACTTTGACGCGCGCGAGGCCAACAAAGCTG 1678
Qy 2059 ttgtggcattgcgaacggagatgcacattcaggagagacacagcgtcctgggagggg 2118
Db 1679 ATCAGGTGCTCAGCCACTGCTCGACACGTTACCGAACTACGAGAGAGTGGGAAGT 1738
Qy 2119 ctcatgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2162
Db 1739 CTCCAGCGCGCGCATGTGCGACAACCTCAGCTGGGACCCAGC 1782

```

RESULT 15

```

US-08-941-445A-6
; Sequence 6, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2542 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 453..2282
; US-08-941-445A-6

```

```

Query Match 4.0%; Score 109.8; DB 3; Length 2542;
Best Local Similarity 50.4%; Pred. No. 1.7e-14;
Matches 323; Conservative 0; Mismatches 312; Indels 6; Gaps 2;

```

```

Qy 1343 aggtgagcagtttaacttttgaagagcagctgtgacagcagcagcagcagcagcagcagcagc 1402
Db 1354 AGGCGAGAGATCAACTGGATGAAGCCCGAATCCTTGAAGCGCGACAGGGTGTCTACCG 1413
Qy 1403 tcactcgggttattcttggaggtgcacaaactgctgaaggtggacaggggctcgaatgagc 1462
Db 1414 TGAGCCCGCTACTAGCGCGAGGAGCTCATCTCCGCGCATCGCAG--GGGATGCGAGCTCG 1470
Qy 1463 tcttaagctccccgaaaaagtgtattggaattgttaattggaattgcaattgcaattgatt 1522
Db 1471 ACAACATCATCGCGCTCACCGGCATCACCGGCATCGGTCAACGGCATGGAGCTGACGAGT 1530
Qy 1523 ggaacccccacacagacaaagtgtct---ccctcatcattattctgtcgatgaactctctg 1579
Db 1531 GGGATCCTAGCAGGACAAAGTACATCATCACCGCAAGTACGACGCAACACCGCAATCCAGG 1590
Qy 1580 gaaagccaaatgtaaaagctgaattgcagaagaggttgggtttaccctgtgaaggagatg 1639
Db 1591 CGAAGCGCTGAACAGGAGGCGTTGCAGGGGAGGCGGTCTTCCGGTCGACAGGAAAA 1650
Qy 1640 ttcctcgtattggctttattggaagactggtattaccagaagagcagcagcagcagcagcagcagc 1699
Db 1651 TCCCACTGATCGCGCTTCATCGCGAGGCTTGAGGAACAGAGGGCCCTGACGTCAATGGCCG 1710
Qy 1700 tggcattccagagctcatgagggagacgtgcaatttgcattgcttggtggtctctggggatc 1759
Db 1711 CCGCCATCCCGAGGCTCATCGAGGAGGACCTCCAGATCGTCTTCTGGGTGACTGGAANA 1770
Qy 1760 caattttgaggtggtgagatctaccagagtcgagttacaagagataaaattccgtggat 1819
Db 1771 AGAAGTTCGAGAAGCTGCTCAAGAGCATGGAGGAAGTATCCGGGCAAGGTGAGGCGG 1830
Qy 1820 ggttggtatttagtttccagatttcccaagaataaactgcaggttgcgatatattgtaa 1879
Db 1831 TGGTGAAGTTCACCGCGCGCTTCTGCTCATCTCATATGCGCGGAGCGGAGTGTCTCGCCG 1890
Qy 1880 tgcacatgagattgaaccttgcggttcttaatacagctatatgtatgcaatatgtgtacag 1939
Db 1891 TCCCGACCGCTTCGAGCCCTGTGGACTCATCAGCTGCGAGGGATGAGATACGGAACGC 1950
Qy 1940 ttcctgtagttcatggaactgggggctcctccgacacagtc 1980
Db 1951 CCTGTGCTTGGCGTCCACCGGTGGGCTCGTGGACACGGGTC 1991

```



Search completed: March 28, 2002, 14:56:11  
Job time: 9785 sec

---



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2002, 13:48:26 ; Search time 3788.66 Seconds  
(without alignments)  
12818.690 Million cell updates/sec

Title: US-09-674-824-1  
Perfect score: 2771  
Sequence: 1 cgcactccactgccttgc.....aaaaaaaaaaaaaaaaaaaaa 2771

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues 34319436

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :		Pending Patents_NA_Main:*	
1:	/cgn2_6/ptodata/1/pna/US0611_COMB.seq:*	17	US-09-345-214-5
2:	/cgn2_6/ptodata/1/pna/US0612_COMB.seq:*	29	US-09-743-980-5
3:	/cgn2_6/ptodata/1/pna/US07_COMB.seq:*	29	US-09-743-980-5
4:	/cgn2_6/ptodata/1/pna/US080_COMB.seq:*	42	US-09-094-436-9
5:	/cgn2_6/ptodata/1/pna/US081_COMB.seq:*	17	US-09-345-214-12
6:	/cgn2_6/ptodata/1/pna/US082_COMB.seq:*	2008	US-09-743-980-12
7:	/cgn2_6/ptodata/1/pna/US083_COMB.seq:*	29	US-09-743-980-12
8:	/cgn2_6/ptodata/1/pna/US084_COMB.seq:*	42	US-09-094-436-12
9:	/cgn2_6/ptodata/1/pna/US085_COMB.seq:*	2008	US-09-094-436-12
10:	/cgn2_6/ptodata/1/pna/US086_COMB.seq:*	48.0	US-09-654-617-385863
11:	/cgn2_6/ptodata/1/pna/US087_COMB.seq:*	2992	US-09-077-564-1
12:	/cgn2_6/ptodata/1/pna/US088_COMB.seq:*	1749	US-09-402-254-52
13:	/cgn2_6/ptodata/1/pna/US089_COMB.seq:*	1752	US-09-625-406-12
14:	/cgn2_6/ptodata/1/pna/US090_COMB.seq:*	24	US-09-625-406-12
15:	/cgn2_6/ptodata/1/pna/US091_COMB.seq:*	25	US-09-654-617-451753
16:	/cgn2_6/ptodata/1/pna/US092_COMB.seq:*	2216	US-09-654-617-451753
17:	/cgn2_6/ptodata/1/pna/US093_COMB.seq:*	2216	US-09-654-617-451753
18:	/cgn2_6/ptodata/1/pna/US094_COMB.seq:*	1620	US-09-625-406-20
19:	/cgn2_6/ptodata/1/pna/US095A_COMB.seq:*	1528	US-09-743-980-6
20:	/cgn2_6/ptodata/1/pna/US095B_COMB.seq:*	17	US-09-345-214-6
21:	/cgn2_6/ptodata/1/pna/US095C_COMB.seq:*	42	US-09-094-436-10
22:	/cgn2_6/ptodata/1/pna/US095D_COMB.seq:*	43.9	US-09-606-304-9
23:	/cgn2_6/ptodata/1/pna/US096A_COMB.seq:*	43.9	US-09-606-304-9
24:	/cgn2_6/ptodata/1/pna/US096B_COMB.seq:*	42.5	US-09-312-544-3744
25:	/cgn2_6/ptodata/1/pna/US096C_COMB.seq:*	1015	US-09-345-214-11
26:	/cgn2_6/ptodata/1/pna/US096D_COMB.seq:*	36.6	US-09-743-980-11
27:	/cgn2_6/ptodata/1/pna/US096E_COMB.seq:*	36.6	US-09-094-436-11
28:	/cgn2_6/ptodata/1/pna/US097A_COMB.seq:*	1015	US-09-606-304-9
29:	/cgn2_6/ptodata/1/pna/US097B_COMB.seq:*	752.8	US-09-654-617-258645
30:	/cgn2_6/ptodata/1/pna/US097C_COMB.seq:*	27.0	US-09-654-617-258645
31:	/cgn2_6/ptodata/1/pna/US098_COMB.seq:*	747	US-09-654-617-258645
32:	/cgn2_6/ptodata/1/pna/US099_COMB.seq:*	712.2	US-09-312-544-395
33:	/cgn2_6/ptodata/1/pna/US099A_COMB.seq:*	25.7	US-09-606-304-3
34:	/cgn2_6/ptodata/1/pna/US099B_COMB.seq:*	648.4	US-09-654-617-269739
35:	/cgn2_6/ptodata/1/pna/US099C_COMB.seq:*	23.4	US-09-654-617-269739
36:	/cgn2_6/ptodata/1/pna/US099D_COMB.seq:*	22.2	US-09-654-617-269739
37:	/cgn2_6/ptodata/1/pna/US099E_COMB.seq:*	614.6	US-09-654-617-269739
38:	/cgn2_6/ptodata/1/pna/US099F_COMB.seq:*	22.2	US-09-654-617-269739
39:	/cgn2_6/ptodata/1/pna/US099G_COMB.seq:*	20.1	US-09-654-617-269739
40:	/cgn2_6/ptodata/1/pna/US099H_COMB.seq:*	555.6	US-09-654-617-269739
41:	/cgn2_6/ptodata/1/pna/US099I_COMB.seq:*	20.1	US-09-654-617-269739
42:	/cgn2_6/ptodata/1/pna/US099J_COMB.seq:*	19.7	US-09-654-617-269739
43:	/cgn2_6/ptodata/1/pna/US099K_COMB.seq:*	547	US-09-654-617-269739
44:	/cgn2_6/ptodata/1/pna/US099L_COMB.seq:*	547	US-09-654-617-269739
45:	/cgn2_6/ptodata/1/pna/US099M_COMB.seq:*	547	US-09-654-617-269739
46:	/cgn2_6/ptodata/1/pna/US099N_COMB.seq:*	547	US-09-654-617-269739
47:	/cgn2_6/ptodata/1/pna/US099O_COMB.seq:*	547	US-09-654-617-269739
48:	/cgn2_6/ptodata/1/pna/US099P_COMB.seq:*	547	US-09-654-617-269739
49:	/cgn2_6/ptodata/1/pna/US099Q_COMB.seq:*	547	US-09-654-617-269739
50:	/cgn2_6/ptodata/1/pna/US099R_COMB.seq:*	547	US-09-654-617-269739
51:	/cgn2_6/ptodata/1/pna/US099S_COMB.seq:*	547	US-09-654-617-269739
52:	/cgn2_6/ptodata/1/pna/US099T_COMB.seq:*	547	US-09-654-617-269739
53:	/cgn2_6/ptodata/1/pna/US099U_COMB.seq:*	547	US-09-654-617-269739
54:	/cgn2_6/ptodata/1/pna/US099V_COMB.seq:*	547	US-09-654-617-269739
55:	/cgn2_6/ptodata/1/pna/US099W_COMB.seq:*	547	US-09-654-617-269739
56:	/cgn2_6/ptodata/1/pna/US099X_COMB.seq:*	547	US-09-654-617-269739
57:	/cgn2_6/ptodata/1/pna/US099Y_COMB.seq:*	547	US-09-654-617-269739
58:	/cgn2_6/ptodata/1/pna/US099Z_COMB.seq:*	547	US-09-654-617-269739
59:	/cgn2_6/ptodata/1/pna/US099AA_COMB.seq:*	547	US-09-654-617-269739
60:	/cgn2_6/ptodata/1/pna/US099AB_COMB.seq:*	547	US-09-654-617-269739
61:	/cgn2_6/ptodata/1/pna/US099AC_COMB.seq:*	547	US-09-654-617-269739
62:	/cgn2_6/ptodata/1/pna/US099AD_COMB.seq:*	547	US-09-654-617-269739
63:	/cgn2_6/ptodata/1/pna/US099AE_COMB.seq:*	547	US-09-654-617-269739
64:	/cgn2_6/ptodata/1/pna/US099AF_COMB.seq:*	547	US-09-654-617-269739
65:	/cgn2_6/ptodata/1/pna/US099AG_COMB.seq:*	547	US-09-654-617-269739

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1392.6	50.3	2491	17	US-09-345-214-5
2	1392.6	50.3	2491	29	US-09-743-980-5
3	1392.6	50.3	2491	42	US-09-094-436-9
4	1392.6	50.2	2008	17	US-09-345-214-12
5	1391.4	50.2	2008	29	US-09-743-980-12
6	1391.4	50.2	2008	42	US-09-094-436-12
7	1329.4	48.0	2606	25	US-09-654-617-385863
8	1329.4	48.0	2606	27	US-09-684-016-385863
9	1327.6	47.9	2992	14	US-09-077-564-1
10	1325.6	47.8	1749	18	US-09-402-254-52
11	1325.6	47.8	1752	24	US-09-625-406-12
12	1268	45.8	2216	25	US-09-654-617-451753
13	1268	45.8	2216	27	US-09-654-617-451753
14	1248.4	45.1	1620	24	US-09-625-406-20
15	1215.8	43.9	1528	17	US-09-345-214-6
16	1215.8	43.9	1528	29	US-09-743-980-6
17	1215.8	43.9	1528	42	US-09-094-436-10
18	1177.8	42.5	1854	64	US-09-312-544-3744
19	1015	36.6	1415	17	US-09-345-214-11
20	1015	36.6	1415	29	US-09-743-980-11
21	1015	36.6	1415	42	US-09-094-436-11
22	752.8	27.2	2360	23	US-09-606-304-9
23	747	27.0	1464	25	US-09-654-617-258645
24	747	27.0	1464	27	US-09-654-617-258645
25	712.2	25.7	1300	64	US-09-312-544-395
26	648.4	23.4	1758	23	US-09-606-304-3
27	614.6	22.2	1034	25	US-09-654-617-269739
28	614.6	22.2	1034	27	US-09-654-617-269739
29	555.6	20.1	619	58	US-09-654-617-269739
30	555.6	20.1	619	58	US-09-654-617-269739
31	547	19.7	555	58	US-09-654-617-269739
32	547	19.7	555	58	US-09-654-617-269739
33	540	19.5	660	53	US-09-200-830-55092
34	454.6	16.4	558	23	US-09-619-643-16489
35	438.2	15.8	555	23	US-09-619-643-16488
36	436.2	15.7	489	58	US-09-654-617-269739
37	436.2	15.7	489	58	US-09-654-617-269739
38	387.8	14.0	430	26	US-09-667-188A-1103
39	387.8	14.0	430	48	US-09-155-008-1836
40	353.6	12.8	469	53	US-09-207-458-39367
41	349	12.6	2478	14	US-09-077-564-3























Db	543	ataagaattatgc	aaatgc	attttacacagaaaaacacattcggattccatgctttggcg	603
Qy	902	gatcacatgaagtgc	accttttttc	atgagatagagacaacgtcgatgggtgttctgcg	961
Db	603	gtgaacatgaagtta	accttctccatgc	atagatagagattcagttgactgggtgtgtttgtg	662
Qy	962	atcatcgtcatata	ccacagaccagga	gttttatatggagataaattttggtcttttgggtg	1021
Db	663	atcatcctcatata	ccagaccctggaa	atttataggagataaagtttgggtcttttgggtg	722
Qy	1022	ataatcagtttcaga	tacacactcctt	gtctatgctgcacgcagagccacctaactcttg	1081
Db	723	ataatcagtttcaga	tacacactcctt	gtctatgctgcacgtgagagctcctttgatctcttg	782
Qy	1082	aattggagagata	tatttatcgacagaa	ttgcattgtttgttgacagattggcatgcga	1141
Db	783	aattggagagata	tatttatcgacagaa	ttgcattgtttgttgcattgatggcatgcga	842
Qy	1142	gccttgcgcagtc	ctcttctgcgc	aaaataatagaccatacagtgattcacagagattccc	1201
Db	843	gtctagtcgcagtc	ctcttctgcgc	aaaataatagaccatacagtgatttataaagactccc	902
Qy	1202	gcagcaccttgtt	tatacataatttag	cacatcaggggtgagcgtcgaagacatactc	1261
Db	903	gcagcattctgt	aaatacataatt	tagcacatcaggggtgagcgtcgaagacatactc	962
Qy	1262	ctgactcggagat	gcctcctgaa	tgtatggagcttttggaaatgggtatttcacagaatggg	1321
Db	963	ctgaccttgggt	tgccactgaa	tgtatggagctcctggagtggttatctccctgaaatggg	1022
Qy	1322	caaggagcgatgc	ctcttgaca	agggtgagggcagttttaaacttttgaaaggagcagttgtga	1381
Db	1023	cgaggagcgatgc	ctcttgaca	agggtgagggcagttttaaacttttgaaagggtcagttgtga	1082
Qy	1382	cagcagatcggat	gtgac	cgctcaggttattctatggaggtgcacaactgctgaag	1441
Db	1083	cagcagatcggaa	tcgtgactgc	aggttatctcgtggaggtgcacaactgctgaag	1142
Qy	1442	gtggacagggc	ctcaatgag	ctcttaagctcccgaaaaagtgtatgaaatggaattgaa	1501
Db	1143	gtggacagggc	ctcaatgag	ctcttaagctcccgaaaaaggtgtatlaaacgggaattgcaa	1202
Qy	1502	atggaattgacat	taattgattgga	ccccccacacagacaagtgtctccctcatcattatt	1561
Db	1203	atggaattgacat	taattgattgga	ccccctgcacacagacaaatgtatccctgtcatcatt	1262
Qy	1562	ctgtcgatgacct	ctctcgtgaa	aggccaaatgtaaagctggaattgcagaaaggagttgggtt	1621
Db	1263	ctgttgatgacct	ctcgtgaa	aggccaaatgtaaagggtgcattgtcagaaaggagctcgggtt	1322
Qy	1622	tacctgtaaagg	ggaggtgttc	ctctgatgtggtcttatggagactggattaccagaaag	1681
Db	1323	tacctgtaaagg	ggctgag	tctctctgatgtggtcttatggaggttggaattacagaaag	1382
Qy	1682	gcattgatctcat	taaaatgc	ctcagagctcatgagggagacgtgcaatttgtca	1741
Db	1383	gcattgatctcat	taaaatgc	ctcagagctcatgagggagagtggttcaatttgtca	1442
Qy	1742	tgcttgga	tctcgtgggatcc	aaatttttgaggctggtgagatctaccaggtcaggttaca	1801
Db	1443	tgcttgga	tctcgtgac	ccagagcttgaagattggaatgcacagatccacagatccca	1502
Qy	1802	aggataaattcc	gtgagtgga	tttccagtttcccaacagataaactgcag	1861
Db	1503	aggataaattcc	gtgagtgga	tttccagtttcccaacagataaactgcag	1562
Qy	1862	gttcgcata	tattgttaa	tgccatgcagatttgaaccttgcgttcttaacagctatatg	1921
Db	1563	gttcgcata	tattgttaa	tgccatccagatttcgaaccttgcgttctcaatcagctatatg	1622
Qy	1922	ctatgcaat	atagttacag	tctctgtagttcatgtgaactcggggcctccgagacacagctcg	1981
Db	1623	ctatgcag	tatggccag	tctctgttgcattgcaactcggggccttagagataccctggtg	1682

Qy	1982	agaaccttaaaccccttttgggtgcacaaaggagagagaggttacaggggtggttcaccgcg	2041
Db	1683	agaacttcaaccccttcgttgagaaatggagagcagggttacaggggtggttcgacccc	1742
Qy	2042	taaccgtggacaagatgttgtggcatttgcgaaccgcgatgtcgacattcaggaggacaca	2101
Db	1743	taaccacagaaaacattgttggacatttgcgaactgcgaatacctacatacacagggaacaca	1802
Qy	2102	agccctcctgggaggggctcatgaaagcgaggagcatgacgaaagaccatactcgtgggaccatg	2161
Db	1803	agtctcctgggaagggttaatgaagcgaggcatgtcaaaagacttcacgtgggaccatg	1862
Qy	2162	cc-ccgagcagtagcagacagatcttcagatgggcttcgtggaccaaacctcagtcagt	2220
Db	1863	ccgctgacaatatcgacaacatcttcagtgggcttcagtcgacccctatgctcagt	1922
Qy	2221	aga 2223	
Db	1923	aaa 1925	
<b>RESULT 6</b>			
US-60-094-436-12			
; Sequence 12, Application US/60094436			
; GENERAL INFORMATION:			
; APPLICANT: Lightner, Jonathan E.			
; APPLICANT: Broglie, Karen E.			
; APPLICANT: E. I. du Pont de Nemours and Company			
; TITLE OF INVENTION: Modification of Starch Biosynthetic Enzyme Gene			
; TITLE OF INVENTION: Expression to Produce Starches in Grain Crops			
; FILE REFERENCE: BB-1147-P1			
; CURRENT APPLICATION NUMBER: US/60/094,436			
; CURRENT FILING DATE: 1998-07-28			
; NUMBER OF SEQ ID NOS: 12			
; SOFTWARE: Microsoft Word Version 7.0A			
; SEQ ID NO 12			
; LENGTH: 2008			
; TYPE: DNA			
; ORGANISM: Zea mays			
US-60-094-436-12			

[illegible]



Db	363	ttggaaagagcaagctcgagctaaagttaacacaaagcattgtctctttgttaacgcgcgaag	422
Qy	722	ctgctccttatgcaaaagtcagggggttggagatgtttgtgggttcgtaccaaatgctc	781
Db	423	ctctccttatgcaagttctgggggtctaggagatgtttgtgttcattgccagttgctc	482
Qy	782	ttgctgtcgttggtcaccgagtgaatgtgttaattgccaagatacttaaatgggtctcctcg	841
Db	483	ttgctgtcgttggtcaccggttgaatgtgttaattgccaagatacttaaatgggtctcctcg	542
Qy	842	ataaaactatgcgaagtcattatcacctgcgaagcacttaagattccactgctttgggg	901
Db	543	ataaagattatgcgaatgcattttacacagaaaaacacattcggatcccatgtttggcg	602
Qy	902	gatacatgaagtgcacctttttcatgagttagagacaaacgtcgatgtgggtgtttgtctg	961
Db	603	gtgaacatgaagtgcacctttccctgagttagagattcgaattgactgggtgtttgttg	662
Qy	962	atcatcgtcatatcacagaccagagaagtttatatggagataaatttgggtgcttttggtg	1021
Db	663	atcatccctcatatcacagaccgtgaaatttatatggagataaatttgggtgcttttggtg	722
Qy	1022	ataatcagttcagatatacacactcctttgctatgctgcagtcgagggccccactaatcctg	1081
Db	723	ataatcagttcagatatacacactcctttgctatgctgcagtcgaggtccctttgatcctg	782
Qy	1082	aattggagatatatttatgacagagattcatgttttgggtgtgaacgattgcatgcga	1141
Db	783	aattggagatatatttatgacagagattgcagtttgggtgtgaacgattgcatgcga	842
Qy	1142	gccttgtgcagtcctcttgcgtgcaaaatagaccatacgggtgtttacagagattccc	1201
Db	843	gtctagtgcagtcctcttgcgtgcaaaatagaccataatgggtgtttataaagactccc	902
Qy	1202	gcagccacttgttatcataataatttagcacatcagggtgttgagacctgcgaagtacatc	1261
Db	903	gcagcattctgttaaatacataatttagcacatcagggtgttgagacctgcgaagcaca	962
Qy	1262	ctgatctgggattgctcctcgaattggtatggagctttagaattgggttatctccagaa	1321
Db	963	ctgaccttgggttgccacctgaattggtatggagctctggagtggttcccigtgaatgg	1022
Qy	1322	caaggaggcatgccccttgacaaggggtgaggcagtttaactctttgaaagggcagctgtga	1381
Db	1023	cgaggaggcatgccccttgacaaggggtgaggcagtttaactctttgaaagggcagctgtga	1082
Qy	1382	cagcaatcggattgtgacgtcagtcagggttatctatcgagaggtcacaaactcgtgaag	1441
Db	1083	cagcagatcgaaatgtgactgtcagtaaggggttatctgtggaggggtcacaaactcgtgaag	1142
Qy	1442	gtggacagggcctcaatgagctcttaagctcccgaaaaagtgtattgaaatggaattgtaa	1501
Db	1143	gtggacagggcctcaatgagctcttaagctcccgaaaaagtgtattgaaatggaattgtaa	1202
Qy	1502	atggaattgacataatgattggaaaccttgcacagacaaatgatcccttgctcatatt	1561
Db	1203	atggaattgacataatgattggaaaccttgcacagacaaatgatcccttgctcatatt	1262
Qy	1562	ctgtcagatgacctcttgaaagggcgaatgtaaagctgaaatgtgcagagaggtgtgggtt	1621
Db	1263	ctgtgtatgacctcttgaaagggcgaatgtaaaggtgcattgcagagagagctgggtt	1322
Qy	1622	tacctgtaaaggaggatgttctctgtattggtccttatctggaagactggattaccagaaag	1681
Db	1323	tacetaaaggcctgatgttctctgatgtgggtcttatctggaaggttggattatcagaag	1382
Qy	1682	gcattgatctcataaaatggccattccagagctcatagaggagagcgtcaaatgtgtca	1741
Db	1383	gcattgatctcataaaatggccattccagagctcatagaggagagcgtcaaatgtgtca	1442
Qy	1742	tgcttggatctggggatcccaatttttgaaggctggatgagatctaccagagtcgagttaca	1801
Db	1443	tgcttggatctgtgtgacccagagcttgaagattggtatgagactcaacagtcgatcttca	1502

[illegible]

Query Match	48.0%;	Score 1329.4;	DB 25;	Length 2606;
Best Local Similarity	80.7%;	Pred. No. 3.5e-113;		
Matches 1606:	Conservative	0: Mismatches 371;	Indels 14;	Gaps 4;

[illegible]







[illegible]

```

RESULT 9
US-03-077-564-1
; Sequence 1, Application US/09077564
; GENERAL INFORMATION:
; APPLICANT: Knight, Mary E.
; APPLICANT: Keeling, Peter L.
; TITLE OF INVENTION: Modification of Starch Synthesis in
; TITLE OF INVENTION: Plants
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Ag Products
;

```







```
QY 1806 taaattccgtggtggttggatttagttgtccagtttcccccagagataaactgcaggttg 1865
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 2108 TAAATTCGTGGATGGTGGATTAGTGTTCAGTTTCCACCGAATAACTGCGGCTG 2167
QY 1866 cgatattgtttaatgccatccagatttgaaccttgcgtgttaatacagctatatctat 1925
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 2168 CGATATATTGTTAATGCCATCCAGATTCGAACCTTGCTGGTCTCAATCAGCTATATGCTAT 2227
QY 1926 gcaatagtacagttccctgttagttcatggaactggggccctccagagacacagtcagac 1985
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 2228 GCAGTATGGACAGTTCCTCTGTGTCTCCATCGAATCGGGGCCCTTAGAGATACCGTGGAGAA 2287
QY 1986 ctccaaccttttgggtgcaaaaggagaggggtacaggggtggcgttctcaccgctaac 2045
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 2288 CTTCACACCTTTCCGTTGAGATGAGAGCGAGGTACAGGGTGGGCATTCGCACCCCTAAC 2347
QY 2046 cgtggacaagatg-ttctggggcatttcgaacccgcgtgtcgacattcaggggagcaaac 2104
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 2348 CACAGAAAACATGTTTGTGGACATTGCGAATTCGAATGCAATATCTACATACAGGGAACACAAGT 2407
QY 2105 cgtctggagggtctcatgaacgcagagcatgaacgaagaccatacgtggaccatgcc- 2163
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 2408 AATAATGGGAAGGCGTAATGAACCCAGGATGTCAAAAGATTCACGTGGGACCATGCCG 2467
QY 2164 ccgagcagtcagcgcagatcttcgagtgggccttcgtgg 2202
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 2468 CTGACAATAACGAACAAATCTTCAGTGGGCCCTTCATCG 2506

RESULT 10
US-09-402-254-52
; Sequence 52, Application US/09402254
; GENERAL INFORMATION:
; APPLICANT: Guan, Hanping
; APPLICANT: Keeling, Peter L.
; TITLE OF INVENTION: PLANT LIKE STARCHES AND THE METHOD OF MAKING THEM IN
; FILE REFERENCE: 2461-52
; CURRENT APPLICATION NUMBER: US/09/402,254
; EARLIER FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: PCT/US98/06660
; EARLIER FILING DATE: 1998-04-03
; EARLIER APPLICATION NUMBER: 60/042,939
; EARLIER FILING DATE: 1997-04-04
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
; LENGTH: 1749
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1749)
US-09-402-254-52
```

```
Query Match 47.8%; Score 1325.6; DB 18; Length 1749;
Best Local Similarity 85.8%; Pred. No. 8.4e-113;
Matches 1509; Conservative 0; Mismatches 239; Indels 10; Gaps 3;

QY 409 taccgtccagctcagcaggagggcccgccgagcccgcccgcccgcccgcccgcccgcccgccg 468
Db 1 tgcgtcgcggagctgagcaggaggggggggggggggggggggggggggggggggggggggggg 60
QY 469 gccccgcgcgtcgtgccagcgtctctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 528
Db 61 gcgccccgcgtcgtgcccggtctctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 117
QY 529 gccccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 588
Db 118 gcatcgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 171
QY 589 ctgaaagggttgctgaggattccatcgacagcataattgtggctgcaagtggagcaggat 648
```

```
Db 172 cctgaagggttgctgaaggttccatcgatacacacagtagttgtggcaagtgagcaagat 231
QY 649 tctgaatcatgagatgcgaatgagcaaacctcaagctaaagttacacgtgacatcgctgttt 708
Db 232 tctgagattggttggaaaggagcagctcagctcaagctcaagctcaagctcaagctcaagctca 291
QY 709 gtgactgggtgaagctgctcctctatgcaaaagtcaggggggtctgggagatgtttgtgttcg 768
Db 292 gtaaacgcggaagcttctccttatgcaagctctgggggtctaggagatgtttgtgttcga 351
QY 769 ttaccaattgctcttctgctgctgctggtccaccagtgatgtgttgaatgccaagatactta 828
Db 352 tfgccagttgctcttctgctgctgctggtccaccgtgtgattgttgaatgcccagatactta 411
QY 829 aatgggtctcttgataaaaaaactatgcaaaaggcattatacacatgcgcgaagcacataagat 888
Db 412 aatgggtctcttgataaaaaaactatgcaaaaggcattatacacatgcgcgaagcacataagat 471
QY 889 ccattgctttggggatcacatgaagtcagcttttttccatgagtatagagacaacgtcgat 948
Db 472 ccattgctttggggatcacatgaagtcagcttttttccatgagtatagagacaacgtcgat 531
QY 949 tgggtgttggctgcatcatccgttcataatcacagaccagggagttttatatagagataatttt 1008
Db 532 tgggtgttggctgcatcatcccttcataatcacagaccctggaaattttatatagagataatttt 591
QY 1009 ggtgctttggtgataatcagttcagatacacacactcctttgctatgctgcagtcgagggcc 1068
Db 592 ggtgctttggtgataatcagttcagatacacacactcctttgctatgctgcagtcgagggct 651
QY 1069 ccactaatccttgaattggaggagatatattatggaacagaattgcattgttgttgaac 1128
Db 652 ccttgcattcctgaattggaggagatatattatggaacagaattgcattgttgttgaac 711
QY 1129 gattgcatgccaagcttgcagctccttctgctgcaaaatatagacatcacgctgtt 1188
Db 712 gattgcatgccaagcttgcagctccttctgctgcaaaatatagacatcacgctgtt 771
QY 1189 tacagagattcccgagcagcacttgtttatacataaatttagcacatcagggtgtggagcct 1248
Db 772 tataaagactcccgagcagcacttgtttatacataaatttagcacatcagggtgtggagcct 831
QY 1249 gcaagtacatactcgtgagtgagtgctcctccgaaatggtatgagcttttagaattgggta 1308
Db 832 gcaagcacatctcgtgagtgagtgctcctccgaaatggtatgagcttttagaattgggta 891
QY 1309 ttccagaaatggcaaggagggcactgccccttgacaagggtgagggcagttacttttgaaa 1368
Db 892 ttccctgaatggcgaggaggagcactgccccttgacaagggtgagggcagttacttttgaaa 951
QY 1369 ggagcagttgtgacagcagatcggttgcacccgtcagtcaggggttattcatggtggaggtc 1428
Db 952 ggtgcagttgtgacagcagatcggttgcacccgtcagtcaggggttattcatggtggaggtc 1011
QY 1429 acaactgctaaagttgagcagggccctcaatgagctcttaagctcccgcaaaaagtgattg 1488
Db 1012 acaactgctaaagttgagcagggccctcaatgagctcttaagctcccgcaaaaagtgattg 1071
QY 1489 aatggaattgtaaatggaattgacatttaattggaaccccccacacagacaagtgctc 1548
Db 1072 aacggaattgtaaatggaattgacatttaattggaaccccccacacagacaagtgctc 1131
QY 1549 cctcatcattattctgctgagctcctctggaaggcccaaatgtaaaagctgaaatcgag 1608
Db 1132 cctctcatattctgctgagctcctctggaaggcccaaatgtaaaagctgaaatcgag 1191
QY 1609 aaggagttgggtttaccctgaaaggagatgttctcctgattggtcttatttgaagagctg 1668
Db 1192 aaggagttgggtttaccctgaaaggagatgttctcctgattggtcttatttgaagagctg 1251
QY 1669 gattaccagaaaggcattgtatctcattaaatggccattccagagctcatgagggagagac 1728
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
```







```

      832 GCAAGCACATCTGACCTGGGTTGCCACCTGAATGGTATGAGCTCTGGAGTGGGA 891
      1309 ttctcagaatggcgaagagcgcacgctccttgacaaaggtgagcagcttaacttttgaaa 1368
      892 TTCCCTGAATGGCGAGAGGCGATGCCCTTGACAAAGGCTGAGCAGCTAATTTTTGAAA 951
      1369 ggaagcagttgtgacagcagatcgagattgtgacccgctcagtcaggttattcatgggaagtc 1428
      952 GGTGCAGTTGTGACAGCAGATCGAATCGTCTGACTGCTCAAGGGTATTCCTGGGAGTCT 1011
      1429 acaactgctgaaggtggacagggcctcaatgagctcttaagctctccgaaaaaagtatttg 1488
      1012 ACAACTGCTGAAGTGGACAGGGCCTCAATGAGCTCTTAAGCTCCAGAAAGAGTGATTA 1071
      1489 aatggaattgtaaatgaaatgacataatgattgaaacccaccacagcaagtgtctc 1548
      1072 AACGGAATGTAAATGAAATGACATTAATGATTGGAAAGCCCTGCCACAGACAAATGTATC 1131
      1549 cctcatcatattctgtcgatgacctctctggaagggccaaatgtaaaagctgaaattgcag 1608
      1132 CCCGTCTATTTCTGTGTGATGACCTCTCTGGAAGGCCAAATGTAAGGTGCATTCGAG 1191
      1609 aaggagttgggtttacctgttaagggaagatgcttctctctgattggctttattggaagatg 1668
      1192 AAGGAGCTGGGTTTACCTATAAGGCCCTCATGTTCTCTGATTGGCTTTATTGGAAGGTG 1251
      1669 gattaccagaaagcattgactctcattaaatgcccattccagagctcatgagggagagac 1728
      1252 GATTATCAGAAAGGCATGTGATCTCATTCACCTTATCATACAGATCTCATGCGGGAAGAT 1311
      1729 gtgcaattgtcatgcttgatctgggataccaaattttgaagctggatgagatctacc 1788
      1312 GTTCAATTTGTCATGCTTTGGATCTGGTGGACCCAGAGCTTGAAGATTGGATGAGATCTACA 1371
      1789 gagtgcagttacaaggataaattccgtggatgggttgattgaggttagtctccagtttccac 1848
      1372 GAGTCGATCTTCAAGGATAAATTTTCGTGGATGGGTGGATTTAGTGTCCAGTTTCCAC 1431
      1849 agaataactgcaggttgcgataatgttaatgcatcgagattgaaaccttgcggtctt 1908
      1432 CGAATAACTGCGGCTGGCATATATTGTTAATGCCATCCAGATTGCAACCTTGTGGTCTC 1491
      1909 aatcagctatatgtatgcaataatgttacagttctctgttagtctcatgaaactgggggctc 1968
      1492 AATCAGCTATATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 1551
      1969 cgagacacagtcgagacctcaacctttgtgtcaaaagagagagaggttacaggtgg 2028
      1552 AGAGATACCGTGGAGAACTTCAACCCCTTCGGTGAGATGGAGAGCAGGGTACAGGGTGG 1611
      2029 gcgttctcaccgctaaacctgggacaagatg-ttgtgggcatctgcgaaccgcgagtctgac 2087
      1612 GCATTCGACCCCTTAACACAGAAACATGTTTGGGACATTCGAACTGCAATATCTAC 1671
      2088 attcaggagcacaaagcgtctctgggaggggctcatgaagcgaggtcatgacgaagacca 2147
      1672 ATACAGGGAACAAAGTCTCTCTGGGAAGGGCTAATGAAGCGAGGCATGTCAAAAGACTT 1731
      2148 tacgtgggaccatgcccc 2165
      1732 CACGTGGGACCATGCCGC 1749

```

RESULT 12

```

US-09-654-617-451753
; Sequence 451753, Application: US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617

```

```

; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 451753
; LENGTH: 2216
; TYPE: DNA
; ORGANISM: Sorghum bicolor
US-09-654-617-451753

```

```

Query Match 45.8%; Score 1268; DB 25; Length 2216;
Best Local Similarity 86.5%; Pred. No. 1.5e-107;
Matches 1422; Conservative 0; Mismatches 220; Indels 2; Gaps 2;

```

```

QY 598 attgctgaggtatcccatcgacagcataaattgttgctgcaagtgcagcagattctgagatc 657
Db 1 attgctgaaaggttcccatcgatgagacagtagttgttgcaagtgcagcagattctgagatc 60
QY 658 atggatgcgaatgagcaaacctcaagctaaagttaacacgtacacgtcgtgttggactgggt 717
Db 61 gtggttggaagagcagcgtcagctaaagttaacacagcattgtctctgcaactggc 120
QY 718 gaagctgctctctatgcaaaagtcagggggttgaggagattgttggtgttctgcttaccat 777
Db 121 gaagcttctctctatgcaaaagtcagggggtctaggagattgttggtgttctgcttaccat 180
QY 778 gctctgtgctgctggttcaccagtgatggttgcataatgcacagatacttaaatgggtcc 837
Db 181 gctctgtgctggttcaccggtgtgattgttgaatgccagataatttaaatgggtacc 240
QY 838 tctgataaaaactatgcaaaagcattatcacatgcagcagacacattaaagtctccatgcttt 897
Db 241 tctgataaagattcacgaaatgcatcttccacagaaaagcattccgattccatgcttt 300
QY 898 gggggatcacatgaagtgcacctttttccatgagtatagagacaacgtcgattgggtgttt 957
Db 301 gggggtgaacatgaagttaacctttttccatgagtacagagatlcagttgactgggtgttt 360
QY 958 gtgcatactccgttcataatcacagaccaggggtttatatatggagataattttgggtgtttt 1017
Db 361 gttgatactccctcattcacagaccctggaaatttatcatggagataaagttgggtgtttt 420
QY 1018 ggtgataatcagttcagatcacacacctcttctgctatgctgcatgcgagcgcccaactaatc 1077
Db 421 ggtgataatcagttcagatcacacacctcttctgctatgctgcatgcgagcctcttggct 480
QY 1078 cttgaaattggagagatatatttatggacagaattgcattgttggatgcacagattggcat 1137
Db 481 cttgaaattggagagatatatttatggacagaattgcattgttggatgcacagattggcat 540
QY 1138 gcaagccttgtgcccagctccttcttgcgcaaaatatagaccatacagctgtttacagagat 1197
Db 541 gcccagtcagtgccagctccttcttgcgcaaaatatagaccatacagctgtttacagagat 600
QY 1198 tcccagcagcctctgttatataataatttagcacatcagggtgtggagcctgcgaagtaca 1257
Db 601 tcccagcagcctctgttaataataatttagcacatcagggtgtaggagcctgcgaagcaca 660
QY 1258 tatcctgatctgggattgcctccttgaaatggatggacttttagaatgggttatttcccgaaa 1317
Db 661 tatcctgatcctgggttgccacctgaaatggatggactctggagtggttattccctgaa 720
QY 1318 tgggcaaggagagcatgccccttgacaaaggggtgagcagtttaacttttgaaaggagcagtt 1377
Db 721 tgggcaaggagagcatgccccttgacaaaggggtgagcagtttaacttttgaaagggtgcagtt 780
QY 1378 gtgacagcagatcgattgtgacccgtcagtcaggggttattcatggagggtcacacactgct 1437
Db 781 gtgacagcagatcgaaattgtgactgtcagtaagggttattcatggagggttcacacactgct 840
QY 1438 gaaggtggacagggcctcaatgagctcttaagctccccgaaaaagctgattgaaaggaaatt 1497
Db 841 gaaggtggacaggggtctcaatgagctcttaagctccccgaaaaagctgattgaaaggaaatt 900

```



QY 1498 gtaaatggaattgacattaatgattggaacccacacagacacagtgctccctcatcat 1557  
 Db 901 gtaaatggaattgacattaatgattggaaccccgacgacacaaatgtatcccttgcat 960  
 QY 1558 tattctgtcatgacacctctctggaaggccaaatgtaaaagctgaattgcgaaggagtg 1617  
 Db 961 tattctgtcatgacacctctctggaaggccaaatgtaaaagctgcatgcsagaaggagctg 1020  
 QY 1618 ggtttacctgaaggaggagtgctctctgtattggtttatttgaagactggattaccag 1677  
 Db 1021 ggtttacctgaaggaggagtgctctctgtattggtttatttgaagactggattaccag 1080  
 QY 1678 aaaggcattgatctatataaaatggccattccagagctcatgaggggagcgtaattt 1737  
 Db 1081 aaaggcattgatctatataaaatggccattccagagctcatgaggggagcgtaattt 1140  
 QY 1738 gtcattgttgatctggggatccaaatttttgaaggctggatgagactaccggagtgagt 1797  
 Db 1141 gtcattgttgatctggggatccaaatttttgaaggctggatgagactaccggagtgagt 1200  
 QY 1798 tacaaggataaaatcccgtagtggttgatttagttccagtttccacagaaataact 1857  
 Db 1201 tacaaggataaaatcccgtagtggttgatttagttccagtttccacagaaataact 1260  
 QY 1858 gcaggttgcgatatattgttaatgccatcgagatttgaaccttgcggtcttaacagcta 1917  
 Db 1261 gcaggttgcgatatattgttaatgccatcgagatttgaaccttgcggtcttaacagcta 1320  
 QY 1918 tatctatgcaatagtgacagttctgttagttcatggaactgggggctcccgagacaca 1977  
 Db 1321 tatctatgcaatagtgacagttctgttagttcatggaactgggggctcccgagacaca 1380  
 QY 1978 gtcgagaccttcaaccttttggtcacaaaggagaggggtacaggtggggcttctca 2037  
 Db 1381 gtcgagaccttcaaccttttggtcacaaaggagaggggtacaggtggggcttctca 1440  
 QY 2038 ccgctaaccgtgacaaagatg-ttggggcattgcgaacgcgagatgcgacattcaggga 2096  
 Db 1441 ccgctaaccgtgacaaagatg-ttggggcattgcgaacgcgagatgcgacattcaggga 1500  
 QY 2097 gcacaagccgtctctggaggggctcatgaagcagggcatgacgaagaccatacagtgga 2156  
 Db 1501 gcacaagccgtctctggaggggctcatgaagcagggcatgacgaagaccatacagtgga 1560  
 QY 2157 ccagtc-ccgagcagtagcagagatctctcagtgggcttcgtggacacacccctacgt 2215  
 Db 1561 ccagtc-ccgagcagtagcagagatctctcagtgggcttcgtggacacacccctacgt 1620  
 QY 2216 catgtagacggggactggggaggt 2239  
 Db 1621 catgtagacggggactggggaggt 1644

RESULT 13

US-09-684-016-451753  
 ; Sequence 451753, Application US/09684016  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Liu, Jingdong  
 ; TITLE OF INVENTION: Annotated Plant Genes  
 ; FILE REFERENCE: 38-21(15097)D  
 ; CURRENT APPLICATION NUMBER: US/09/684,016  
 ; CURRENT FILING DATE: 2000-10-10  
 ; PRIOR APPLICATION NUMBER: US 09/654,617  
 ; PRIOR FILING DATE: 2000-09-05  
 ; NUMBER OF SEQ ID NOS: 463173  
 ; SEQ ID NO 451753  
 ; LENGTH: 2216  
 ; TYPE: DNA  
 ; ORGANISM: Sorghum bicolor  
 US-09-684-016-451753

Query Match 45.8%; Score 1268; DB 27; Length 2216;  
 Best Local Similarity 86.5%; Pred. No. 1.5e-107;  
 Matches 1422; Conservative 0; Mismatches 220; Indels 2; Gaps 2;

QY 598 atttctgagagattcccatcgacagcataaattgtgctgcgaagtgcagagattcttgagatc 657  
 Db 1 atttctgagagattcccatcgacagcataaattgtgctgcgaagtgcagagattcttgagatc 60  
 QY 658 atggatgcgaatgagcaaacctcaagctaaagtacacgtagcatcgtgtttgtgaactggt 717  
 Db 61 gtggttggaaaggagcaagctgcagctaaagtacacaaagcattgtctttgtaactggc 120  
 QY 718 gaagctgctccttatcaaaagtcagggggttggagatgtttgtgttcgttttaaccaatt 777  
 Db 121 gaagctgctccttatcaaaagtcagggggttggagatgtttgtgttcgttttaaccaatt 180  
 QY 778 gctcttgcgtcgtgtgcacagtgatggttgtaatgccaaagatacttaaaatgggtcc 837  
 Db 181 gctcttgcgtcgtgtgcacagtgatggttgtaatgccaaagatacttaaaatgggtcc 240  
 QY 838 tctgataaaaaactatgcgaaggcattatacaactgcgaagcacacattaaagattccatt 897  
 Db 241 tctgataaaaaactatgcgaaggcattatacaactgcgaagcacacattaaagattccatt 300  
 QY 898 gggggatcacatgaagtgccttttttcatgatatagagacaacgtcgattgggtgttt 957  
 Db 301 gggggatcacatgaagtgccttttttcatgatatagagacaacgtcgattgggtgttt 360  
 QY 958 gtcgatactccgtcatatcacagaccaggaagtttatatggagataaattttgggtgttt 1017  
 Db 361 gtcgatactccgtcatatcacagaccaggaagtttatatggagataaattttgggtgttt 420  
 QY 1018 ggtgataactcagttcagatacacacactccttttgcatactgcagagccccaactaact 1077  
 Db 421 ggtgataactcagttcagatacacacactccttttgcatactgcagagccccaactaact 480  
 QY 1078 cttgtaattgggagagataattttatgacagaaatttgatgttggatgacattggcat 1137  
 Db 481 cttgtaattgggagagataattttatgacagaaatttgatgttggatgacattggcat 540  
 QY 1138 gccagccttgcagcagtccttcttgcgcacaaatatagaccatacgggttttacagagat 1197  
 Db 541 gccagccttgcagcagtccttcttgcgcacaaatatagaccatacgggttttacagagat 600  
 QY 1198 tcccgacagacccctgtttatataataatttagcacatcaggggtgcgagcctgcagatac 1257  
 Db 601 tcccgacagacccctgtttatataataatttagcacatcaggggtgcgagcctgcagatac 660  
 QY 1258 tatctctgattgggattgctcctcctgaattggtatgagcctttagaatgggtattccagaa 1317  
 Db 661 tatctctgattgggattgctcctcctgaattggtatgagcctttagaatgggtattccagaa 720  
 QY 1318 tgggcaaggagagcagtccttgcacaaagggtgcaggcagtttaactttttgaaaggagcatt 1377  
 Db 721 tgggcaaggagagcagtccttgcacaaagggtgcaggcagtttaactttttgaaaggagcatt 780  
 QY 1378 gtgacagcagatcgagattgtgacctcagtcaggtttatctatggagagtcacaactgt 1437  
 Db 781 gtgacagcagatcgagattgtgacctcagtcaggtttatctatggagagtcacaactgt 840  
 QY 1438 gaagtgacagaggcctcaatgagctcttaagctcccgaaaaagtgtattgaatggaatt 1497  
 Db 841 gaagtgacagaggcctcaatgagctcttaagctcccgaaaaagtgtattgaatggaatt 900  
 QY 1498 gtaaatggaattgacataatgattggaaacccacacagacagaaagtgtctcctccatcat 1557  
 Db 901 gtaaatggaattgacataatgattggaaacccctgcagcgggacaaaatgtatcctctgtcat 960  
 QY 1558 tattctgtcagtagcctctctggaaaggccaaatgtlaaagctgaattgcagagaggttg 1617  
 Db 961 tattctgtcagtagcctctctggaaaggccaaatgtlaaagctgaattgcagagaggttg 1020  
 QY 1618 ggtttacctgtaaggaggatgttctctctgattgttctctgtggtttattggaagcag 1677



||||| 1021 gggttaacctaaaggccctgaagttccctctgattggttttatttggaagattggattacag 1080  
Qy 1678 aaaggcattatctcattaaatggccattccagagctcatgagggagcgctgcaattt 1737  
Db 1081 aaaggcattatctcattaaatggccattccagagctcatgagggagcgctgcaattt 1140  
Qy 1738 gtcattgttgatctggggtaccatttttgaaggctggatgagatctaccgagctgagt 1797  
Db 1141 gtcattgttgatctggggtaccatttttgaaggctggatgagatctaccgagctgagt 1200  
Qy 1798 tacaagataaaattccctggatgggttgatttagttccagtttcccccacagaataact 1857  
Db 1201 tcaagataaaattccctggatgggttgatttagttccagtttcccccacagaataact 1260  
Qy 1858 gcaggttgcgatatattgttaagtcacatcgagatttgaaccttgcggtcttaatacagcta 1917  
Db 1261 gcggctgcgatatattgttgaagtcacatcgagatttgaaccttgcggtcttaatacagcta 1320  
Qy 1918 tatgtatgaatatggtacagttctctgttagttcattgaagctggggccctcccgagacaca 1977  
Db 1321 tatgtatgagatggtacagttctctgttagttcattgaagctggggccctcccgagacaca 1380  
Qy 1978 gtcagaccttcaaccttttgggtgcaaaaggagagaggtacaggggtggggtctctca 2037  
Db 1381 gtcagaccttcaaccttttgggtgcaaaaggagagaggtacaggggtggggtctctca 1440  
Qy 2038 ccgtaaccgtggagacagatg-tctgtggcattgacacccgagatgacacattcagggga 2096  
Db 1441 cccttaaccgggaaacagatg-tctgtggcattgacacccgagatgacacattcagggga 1500  
Qy 2097 gcaaaagccctctctggaggggtctatgaagcgagggcatgacgaaagacacatcagtgga 2156  
Db 1501 gcaaaagccctctctggaggggtctatgaagcgagggcatgacgaaagacacatcagtgga 1560  
Qy 2157 ccattgcc-cgagagcattacagacagatctctcagtgagggcctctgagacacacccctcgt 2215  
Db 1561 ccattgcc-cgagagcattacagacagatctctcagtgagggcctctgagacacacccctcgt 1620  
Qy 2216 catgtagacggggagctgggaggt 2239  
Db 1621 catgtaaaggggacacacagtggt 1644

## RESULT 14

US-09-625-406-20  
Sequence 20, Application US/09625406  
GENERAL INFORMATION:  
APPLICANT: Keeling, Peter  
TITLE OF INVENTION: Starch Encapsulation  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/625,406  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/941,445  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Winner, Ellen P

REGISTRATION NUMBER: 28,547  
REFERENCE/DOCKET NUMBER: 89-97  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1620 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1620  
US-09-625-406-20

Query Match 45.1%; Score 1248.4; DB 24; Length 1620;  
Best Local Similarity 86.2%; Pred. No. 9.8e-106;  
Matches 1392; Conservative 0; Mismatches 221; Indels 1; Gaps 1;

Qy 553 gacgcgcgcgtgggggaactcgcgcgaacctctcgtcgaaggattgctgaggattcc 612  
Db 4 gtcgcgcgcgtgggggaactcgcgcgaacctctcgtcgaaggattgctgaggattcc 63  
Qy 613 atcgacgacgataaattgtggctgcaagtgcagcaggttctgagatcatggatgcgaatgag 672  
Db 64 atcgatnaacacagtagttgtggcaagtgacgaagattctgagatttgggttggaaaggag 123  
Qy 673 caacctcaagctaaagtacacgttagcatcgtgtttgtgactggtgaaagctgctccttat 732  
Db 124 caagctcgagctaaagtgcgaagtgacgaagatttgcgttttaacccggcgaagcttcccttat 183  
Qy 733 gcaaaagtcaggggggtggagatgttctggttcgttaccatctctctgctgcgtcgt 792  
Db 184 gcaaaagtcaggggggtggagatgttctggttcgttaccatctctctgctgcgtcgt 243  
Qy 793 ggtcaccgagtgatggtgttaagtcgaagatacttaaatgggtcctctgataaaaaactat 852  
Db 244 ggtcaccgagtgatggtgttaagtcgaagatacttaaatgggtcctctgataaaaaactat 303  
Qy 853 gcaaaagtcatactacactgcgaagacacatcaagattccatcgttgggggacacacgaa 912  
Db 304 gcaaaagtcatactacactgcgaagacacacatcaagattccatcgttgggggacacacgaa 363  
Qy 913 gtagcctttttcatgagtagagacacacgctgaggtggttctgctgacatcccgta 972  
Db 364 gttacccttttccatgagtagatgagattcagttgacgtgggtgttgggtgattcattccctca 423  
Qy 973 tatcacagacaggaagtttatatggagataattttgggtgcttttgggtgataaatcagttc 1032  
Db 424 tatcacagacagctggaaatttatatggagataattttgggtgcttttgggtgataaatcagttc 483  
Qy 1033 agatacacactccttctgctatgctgcgagggcccccactaatccttgaattgggagga 1092  
Db 484 agatacacactccttctgctatgctgcgagggcccccactaatccttgaattgggagga 543  
Qy 1093 tatatttatggacagaattgcatgttttggtaacagattggcagccttgggtgacga 1152  
Db 544 tatatttatggacagaattgcatgttttggtaacagattggcagccttgggtgacga 603  
Qy 1153 gtccttcttctgcacaaatagaccatacgggtgttttaccagagattccccgagacacccctt 1212  
Db 604 gtccttcttctgcacaaatagaccatacgggtgttttaccagagattccccgagacacccctt 663  
Qy 1213 gttatacataatttagcacatcaggggtgtagagcctcaggtacatctcctgattcgtgga 1272  
Db 664 gtaatacataatttagcacatcaggggtgtagagcctcaggtacatctcctgattcgtgga 723  
Qy 1273 ttgcctcctgaatggtatggagccttttagaattgggtatttccagaaatgggcaagagagcat 1332



Db 724 TTGCCACCTGAATGGTATGAGCTCTGGAGTGGGTATTCCTCGAATGGCGGAGGCGCAT 783  
QY 1333 gcccttgacaaaggtgagcagtaactttttgaaaggagcagttgtgacagcagatcgg 1392  
Db 784 GCCCTTGACAAGGGTGAGGCGAGTAAATTTTGAAGGGTGACAGTGTGTGACAGCATCGA 843  
QY 1393 atttgaccctcagtcaggttattcatggagggtcacactgctgaaggtggacagggc 1452  
Db 844 ATCGTGACTGTCAGTAAGGGTATTCGTGGGAGGTCACAACCTGCTGAAGGTGGACAGGC 903  
QY 1453 ctoaatgagctcttaagctcccgaaagagtgattgaaatggaattgaaatgac 1512  
Db 904 CTCGAATGAGCTCTTAAGCTCCAGAAAGAGTATTAAACGGAATTGTAATGGAATTGAC 963  
QY 1513 attaatgattgaaccocccacacaaagtgatctccctcatcattattctctgaterac 1572  
Db 964 ATTAATGATTGAACCCCTCCACAGACAATGATATCCCTGTCAATATTCTGTGTGATGAC 1023  
QY 1573 ctctctggaagggccaaatgtaagctgaattgcagaagaggttggtttaccctgtaagg 1632  
Db 1024 CTCTCTGGAAGGCGCAAAATGTAAGGTGCATTCGACAAGSAGCTGGGTTTACCTATAAGS 1083  
QY 1633 gaggatgtctctctgattgcttattggaagactggattaccagaaagccattgactc 1692  
Db 1084 CCTGATGTTCTCTGATTGGCTTTATTGGAAGGTTGGATTATCAGAAAGGCATTGATCTC 1143  
QY 1693 attaaatggccattccagactcatgaggagagctgaatttctgactgctgactc 1752  
Db 1144 ATTCAACTTATACACATCTCATCGGGAAGATGTTCAATTTGCTATGCTGTGATCT 1203  
QY 1753 gggatgccaatatttgaagcgtggatgagatctaccagctcaggttcacaggaataatc 1812  
Db 1204 GGTGACCCAGAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCAAGGATAAATT 1263  
QY 1813 cgtgagatgggttgattgattgctcagttccagttccacagaaataactcaggttgcgata 1872  
Db 1264 CGTGGATGGGTGGATTTAGTGTTCAGTTTCCACCGAATAACTCCCGGCTCCGATATA 1323  
QY 1873 ttgttaatgcatcgagattgaacctgcggtcttaacagctatatctcattgcaatat 1932  
Db 1324 TTGTTAATGCCATCCAGATTCGAACCTTGTGGTCTCAATCAGCTATATGCTATGAGTAT 1383  
QY 1933 ggtacagttctctgattgctgaactggggcctcccgagacacagctcagacacctcaac 1992  
Db 1384 GGCACAGTTCCTGTTGCTCATCACTGGGGGCTTAGAGATACCCGTGGAGAACTTCAAC 1443  
QY 1993 ccttttgggtcaaaagagagaggttacaggtggcggtctccacgcgtacacgtggac 2052  
Db 1444 CCTTCGGTGAGAAATGGAGAGCAGGTACAGGGTGGGCTTCGCACCCCTTAACACAGAA 1503  
QY 2053 aagatg-ttgtggcattgcgaaccgcatgctgacattcagggagcagacaaagccgtcctg 2111  
Db 1504 AACATGTTGTGACATTCGGAATGCATATCTACATACAGGGAACAAAGTCCCTCTG 1563  
QY 2112 ggaagggctaatgaagcggagcgtacgaaagaccatacgtgggacatgcccc 2165  
Db 1564 GGAAGGCTAATGAAGCGGCGCATGCAAAAGACTTCACGTGGGACCATGCCGC 1617

RESULT 15

US-09-345-214-6/c  
; Sequence 6, Application US/09345214  
; GENERAL INFORMATION:  
; APPLICANT: Broglie, Karen E.  
; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE  
; FILE REFERENCE: BB-1147  
; CURRENT APPLICATION NUMBER: US/09/345,214  
; CURRENT FILING DATE: 1999-06-30  
; EARLIER APPLICATION NUMBER: 060/094,436  
; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 6  
; LENGTH: 1528  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-345-214-6  
  
Query Match 43.9%; Score 1215.8; DB 17; Length 1528;  
Best Local Similarity 87.7%; Pred. No. 9.4e-103;  
Matches 1339; Conservative 0; Mismatches 187; Indels 1; Gaps 1;  
  
QY 678 tcaagctaaagtacacgtagcatcgtttgtgactggtgaagctgctcccttatgcaaa 737  
Db 1528 TCGAGGCTAAAGTAACACAAAAGCATTTGCTTTGTAAACCGGGAAGCTTCTCCTTATGCAAA 1469  
QY 738 gtcagggggggttggagagatggttgggttcggttaccaaatgctctctgctgctgctggtca 797  
Db 1468 GTCTGGGGTCTAGGAGATGTTTGTGGTTCATTTGCCAGTGTCTTTCGTCTCGTGGTCA 1409  
QY 798 ccgagtgatggttgaatcccaagatacttaaatgggtcctctgataaaaaactatgcaaa 857  
Db 1408 CCGTGTGATGGTGTAAATGCCAGATATTTAAATGTACTCCGATPAAGAAATATGCAAA 1349  
QY 858 ggcattatcacactgcgaagcacatgaattccatgcttttgggggacatcacatgaagtac 917  
Db 1348 TGCATTTTACACAGAAAACACATTCGGATTCATGCTTTTGGGGTGAACATGAAGTTAC 1289  
QY 918 ctttttcataggtatagagacaacgtcagattgggttgggttctgctgctcctcctcctc 977  
Db 1288 CTTCTTCCATGATATAGAGATTCAGTTGACTGGGTGTTTGTGTGATCATCCCTCATATCA 1229  
QY 978 cagaccagaaagtattataggagataatttttgggtccttttgggtgataatcagttcagata 1037  
Db 1228 CAGACCTGGAATTTATATGGAGATAAAGTTTGGTGTCTTTTGGTGAATACAGTTTCAGATA 1169  
QY 1038 cacactcctttgctatgctgcatgcgagggcccaactaaaccttgaattggggaggatat 1097  
Db 1168 CACACTCCTTTGCTATGCTGCATGTGAGGCTCCTTTTGATCCTTGAATTTGGAGGATATAT 1109  
QY 1098 ttatgagacaaattgctattgttggatgaacgattggcattggcagcttgcagctcct 1157  
Db 1108 TTATGACAGAAATGTGATGTTTGTCTCAATGATTGGCATGCCAGTCTAGTGCCAGTCT 1049  
QY 1158 tcttgctcaaaaatagaccatacgggttttacagagattcccgagcagcccttcttat 1217  
Db 1048 TCTTGCTGCAAAATATAGACCATATGTTGTTTATAAAGACTCCCGCAGCATCTTCTGAAT 989  
QY 1218 acataatttagcacatcaggggtgtggagcctgcaagtacatatcctgctgctggtatgcc 1277  
Db 988 ACATAATTTAGCACATCAGGGTGTAGAGCTGCAAGCACATATCCTTGACCTTGGTGGCC 929  
QY 1278 tccgtaaggtatggagctttgaaatgggtatttccagaatggcgaagagcagctccct 1337  
Db 928 ACCTGAATGGTATGGAGCTCTGGAGTGGGTATTTCCCTGAAATGGCGGAGGAGGATGCCCT 869  
QY 1338 tgacaaggggtgagcgagttlaactttttgaaaggagcagttgtgacagcagatcggaattgt 1397  
Db 868 TGACAAGGGTGGAGCATTAATTTTGAAGGTGTCAGTTGTGCACAGCAGATCAATCGT 809  
QY 1398 gaccctcagtcaggggttatcatcagggaggttcacaactgctgaaggtggacagggcctcaa 1457  
Db 808 GACTGTCAAGGGTATTTCTGGGAGGTCACAACCTGCTGAAAGGTGGACAGGCGCTCAA 749  
QY 1458 tgagctcttaagctcccgaaaaagtattgaatgaaattgtaaatggaattgaacattaa 1517  
Db 748 TGAGCTCTTAAAGCTCCAGAAAAGAGTGTATTAAACGGAAATTTGAAATGGAATGACATTAA 689  
QY 1518 tgattggaaccccccacacagacagtaagtgtctccctcattcatttctgtcgtatgacctctc 1577  
Db 688 TGATTGGAACCCCTGCCACACAGACAAATGTATCCCTGTCAATTATCTGTTGATGACCTCTC 629  
QY 1578 tggaaaggccaaatgtaaagctgaattgcagaaggaggttgggtttaccctgtaaggagga 1637



```

|||||
Db 628 TGGAAAGGCCAATGTAAGGTGCTATTCAGAGAGAGCTGGGTTTACCTATAAGGCCTGA 569
Qy 1638 tgttccctctgattggtcttattggaagacttggtattaccagaaagcattgatctcattaa 1697
Db 568 TGTTCCTCTGATTGGCTTTATTGGAAGGTTGGATTATCAGAAAGGCATTGATCTCATTTCA 509
Qy 1698 aatggccattccagagatcattgaaggaggacgctgcaatttgcattgcttgatctgggga 1757
Db 508 ACTTATCATACCAAGATCTCATGCGGGAAGATGTTCAATTTGTCATGCTGATCTGGTGA 449
Qy 1758 tceaattttgaagctggatgagatctaccgagtcgagttacaagataaaattccgtgg 1817
Db 448 CCCAGAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCAAGGATAAATTTCTGG 389
Qy 1818 atgggttggatttagtgttccagtttccacagaaataactgcaggttgcgatatattgtt 1877
Db 388 ATGGGTTGGATTAGTCTTCCAGTTTCCCAACCGAATAACTCCGGCTGCCGATATATTGTT 329
Qy 1878 aatgceatcgagatttgaaccttgcggtcttaatacagctatgctatgcaaatatggtac 1937
Db 328 AATGCCATCCAGATTCCGAACCTTTGTGCTCTCAATCAGCTATATGCTATGAGTATGGCAC 269
Qy 1938 agttccctgtagttcatggaactggggcctccgagacacagtcgagaccttcaaccttt 1997
Db 268 AGTTCTCTGTTGTCATGCAACTGGGGGCCCTTAGAGATACCGTGGAGAACTTCAACCCTTT 209
Qy 1998 tgggtcaaaaggagagaggtacaggggtggcggttctcaccgcctaaccctggacaagat 2057
Db 208 CGGTGAGAATGGAGAGCAGGGTACAGGGTGGGCTTCCGCCCTTAACCAACAGAAACAT 149
Qy 2058 gttgtgggcaatttcgaaaccgagatgctcgacattcaggaggagcacaagccgtccctgggagg 2117
Db 148 GTTGTGGACATTGGCAACTGCAATATCTACATACAGGGAAACACAAGTCTCTCTGGGAAG 89
Qy 2118 gctcatgaagcgggcatgacgaagaccatacgtgggaccatgcc-ccgagcagtacga 2176
Db 88 GCTAATGAAGCGAGGCATGTCAAAAGACTTCACGTGGGACCATGCCGCTGAACAATACGA 29
Qy 2177 gcagatcttcgagtgggccttcgtgga 2203
Db 28 ACAAACTTCCAGTGGGGCTTTCATCGA 2

```

Search completed: March 28, 2002, 16:19:36  
Job time: 9070 sec



Result No.	Query Match	Score	Length	DB	ID	Description
1	100.0	2771	5	2771	5	US-09-674-824-1
2	91.7	2541.8	8	2862	8	US-09-508-377-11
3	80.8	2239	7	2239	7	US-09-552-677-1
4	50.3	1393	5	2383	7	US-09-931-297-1
5	20.5	567.6	20	10337	8	US-09-508-377-13
6	19.5	540	5	660	8	US-09-873-402A-83353
7	14.4	399	8	589	8	US-09-873-402A-67388
8	12.8	355.2	7	561	7	US-09-865-439A-118661
9	11.1	307.8	41	407	6	US-09-985-678-88686
10	11.1	307.8	41	411	6	US-09-985-678-83476
11	11.1	307.6	9	598	9	US-10-021-323-969
12	10.8	299	8	5072	8	US-09-508-377-15
13	10.6	293.6	6	412	7	US-09-865-439A-21857
14	10.6	293.6	7	635	7	US-09-636-664A-17004
15	10.3	286.6	5	577	7	US-09-849-529A-17365
16	10.1	279	6	335	6	US-09-985-678-9898
17	9.8	272.4	9	2418	9	US-10-044-543-25
18	9.7	269.4	9	372	7	US-09-865-439A-29137
19	9.2	255.8	6	316	6	US-09-985-678-39564
20	9.2	255.8	6	2348	9	US-10-044-543-5
21	9.1	251.6	9	951	5	US-09-819-091A-6343
22	9.0	249.2	6	286	6	US-09-985-678-41347
23	8.6	237.4	6	289	6	US-09-985-678-90311



QY 241 cctctccgctctccctctcccgccgcccacacccatgcccgcgagcgctggc 300  
DB 241 cctctccgctctccctctcccgccgcccacacccatgcccgcgagcgctggc 300  
QY 301 gcgggtgctctgccccagctccgctcgccgagatcccgagcgcccgcccgccg 360  
DB 301 gcgggtgctctgccccagctccgctcgccgagatcccgagcgcccgcccgccg 360  
QY 361 tccgctgctctgccccagctccgctcgccgagatcccgagcgcccgcccgccg 420  
DB 361 tccgctgctctgccccagctccgctcgccgagatcccgagcgcccgcccgccg 420  
QY 421 ctaagcaggagggccccgcgcccgcgcccgcgagcagcagcagcagcagcagc 480  
DB 421 ctaagcaggagggccccgcgcccgcgcccgcgagcagcagcagcagcagcagc 480  
QY 481 gtgcaggctctccgcgcccgcgcccgcgcccgcgcccgcgcccgcgcccgcg 540  
DB 481 gtgcaggctctccgcgcccgcgcccgcgcccgcgcccgcgcccgcgcccgcg 540  
QY 541 ccgcccctgcgagcgcccgcgcccgcgcccgcgcccgcgcccgcgcccgcg 600  
DB 541 ccgcccctgcgagcgcccgcgcccgcgcccgcgcccgcgcccgcgcccgcg 600  
QY 601 gctgagattccatcgacagataattgtggctgcaagtgcgagcagcagcagcag 660  
DB 601 gctgagattccatcgacagataattgtggctgcaagtgcgagcagcagcagcag 660  
QY 661 gatcggaatgagcaactcaagctaaagttaacgctgacgctgctgtgtgtgactgg 720  
DB 661 gatcggaatgagcaactcaagctaaagttaacgctgacgctgctgtgtgtgactgg 720  
QY 721 gctgctctctatgcaagctgaggggttgaggagatgtgtgtgtgtgtgtgtgtgt 780  
DB 721 gctgctctctatgcaagctgaggggttgaggagatgtgtgtgtgtgtgtgtgtgt 780  
QY 781 ctgtgctgctgtgtacacgagtgatggtgttaatgcaagatacttaaatgggtctct 840  
DB 781 ctgtgctgctgtgtacacgagtgatggtgttaatgcaagatacttaaatgggtctct 840  
QY 841 gataaaactatgcaagcattatcacctgcgaagcattagatccatgcttggg 900  
DB 841 gataaaactatgcaagcattatcacctgcgaagcattagatccatgcttggg 900  
QY 901 ggatcacatgaagtgcactttttctatgagtatagagacaacgctgattgggtgtgt 960  
DB 901 ggatcacatgaagtgcactttttctatgagtatagagacaacgctgattgggtgtgt 960  
QY 961 gatctcgtcatatcacagaccaggaagtattatgagataatttgggtgtgtgtgtgt 1020  
DB 961 gatctcgtcatatcacagaccaggaagtattatgagataatttgggtgtgtgtgtgt 1020  
QY 1021 gataatcagttcagatacacactcttgcctatgctgacgagcccccactaatcctt 1080  
DB 1021 gataatcagttcagatacacactcttgcctatgctgacgagcccccactaatcctt 1080  
QY 1081 gaattggaggagataattattgagcagaattgcatgtgtgtgtgaaacgattggcagtc 1140  
DB 1081 gaattggaggagataattattgagcagaattgcatgtgtgtgtgaaacgattggcagtc 1140  
QY 1141 agcctgtgacagctctcttctgctgcaaaatagacacatcagctgtgttaccagatttcc 1200  
DB 1141 agcctgtgacagctctcttctgctgcaaaatagacacatcagctgtgttaccagatttcc 1200  
QY 1201 cgcagcaccttctgtatataataatttagcacatcagggtgtggagcctgcaagtacatat 1260  
DB 1201 cgcagcaccttctgtatataataatttagcacatcagggtgtggagcctgcaagtacatat 1260  
QY 1261 cctgattctgggatttgcctcctgaaatggatgagcttttagaattgggtatttccagaatgg 1320  
DB 1261 cctgattctgggatttgcctcctgaaatggatgagcttttagaattgggtatttccagaatgg 1320  
QY 1321 gcaaggagcgctcccttgacaaagggtgaggcagcttaactttttgaaaggagcagttgtg 1380

DB 1321 gcaaggagcgctcccttgacaaagggtgaggcagcttaactttttgaaaggagcagttgtg 1380  
QY 1381 acagcagatcgagatgtgacccgtcagtggtttattcattggagggtcacaaactcgtgaa 1440  
DB 1381 acagcagatcgagatgtgacccgtcagtggtttattcattggagggtcacaaactcgtgaa 1440  
QY 1441 ggtggacagggcctcaatgagctcttaagctcccccagaaagtgtattgaatggaattgta 1500  
DB 1441 ggtggacagggcctcaatgagctcttaagctcccccagaaagtgtattgaatggaattgta 1500  
QY 1501 aatgggaattgacatttaattggaacccccacacagaaagtgtctccctcatcattat 1560  
DB 1501 aatgggaattgacatttaattggaacccccacacagaaagtgtctccctcatcattat 1560  
QY 1561 tctgtcagatgaactctctggaagggccaaatgttaaagctgaaatgcagaaaggagtgggt 1620  
DB 1561 tctgtcagatgaactctctggaagggccaaatgttaaagctgaaatgcagaaaggagtgggt 1620  
QY 1621 ttacctgaaggagagatgttctctctgattggctttatttgggaagcagctgattaccagaaa 1680  
DB 1621 ttacctgaaggagagatgttctctctgattggctttatttgggaagcagctgattaccagaaa 1680  
QY 1681 ggcattgactcattaaaaatggccattccagagctcatgaggagagcagcgaatttgc 1740  
DB 1681 ggcattgactcattaaaaatggccattccagagctcatgaggagagcagcgaatttgc 1740  
QY 1741 atgcttgagatctggggtccaaatgttgaagcgtgagatgcagctaccagctcagagttac 1800  
DB 1741 atgcttgagatctggggtccaaatgttgaagcgtgagatgcagctaccagctcagagttac 1800  
QY 1801 aaggataaaatcccgctggatgggttgatttagtctccagtttccacagaaataactgca 1860  
DB 1801 aaggataaaatcccgctggatgggttgatttagtctccagtttccacagaaataactgca 1860  
QY 1861 ggttgagatataattgtaagcagcagatgtgaacccctgggttcttaactcagctatat 1920  
DB 1861 ggttgagatataattgtaagcagcagatgtgaacccctgggttcttaactcagctatat 1920  
QY 1921 gctatgcaatattggtacagttctctgttagttcaatgggaactgggggctccgagacagtc 1980  
DB 1921 gctatgcaatattggtacagttctctgttagttcaatgggaactgggggctccgagacagtc 1980  
QY 1981 gagacctcaaccccttttgcgaagggagagaggttacaggggtggggttctccacgc 2040  
DB 1981 gagacctcaaccccttttgcgaagggagagaggttacaggggtggggttctccacgc 2040  
QY 2041 ctaacccgtgacaaagtgtgtgggcatgtgcaacccgcagatgtcgacattcaggggagcac 2100  
DB 2041 ctaacccgtgacaaagtgtgtgggcatgtgcaacccgcagatgtcgacattcaggggagcac 2100  
QY 2101 aagccgtctctgggaggggctcatgaaagcagcagcagcagcagcagcagcagcagcagc 2160  
DB 2101 aagccgtctctgggaggggctcatgaaagcagcagcagcagcagcagcagcagcagcagc 2160  
QY 2161 gcccccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2220  
DB 2161 gcccccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2220  
QY 2221 agacggggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2280  
DB 2221 agacggggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2280  
QY 2281 tcccttcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2340  
DB 2281 tcccttcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2340  
QY 2341 caatgcgctgctgt 2400  
DB 2341 caatgcgctgctgt 2400  
QY 2401 gcggtgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2460



Db 2401 gcggtgacagcttcgggtgacgacagattacagattttggggaataaaggagggatgtgct 2460  
QY 2461 gcaggatggttaacagcaaaagcaccactcagatggcagcctctctgtccgtgttacagct 2520  
Db 2461 gcaggatggttaacagcaaaagcaccactcagatggcagcctctctgtccgtgttacagct 2520  
QY 2521 gaaatcagaacaaactggtagctcttttagccttagtgatgtggaagtgttggctctct 2580  
Db 2521 gaaatcagaacaaactggtagctcttttagccttagtgatgtggaagtgttggctctct 2580  
QY 2581 gtgtatgttctgtcttagctgacaaatatttgcctgttggagaaatttatctttg 2640  
Db 2581 gtgtatgttctgtcttagctgacaaatatttgcctgttggagaaatttatctttg 2640  
QY 2641 ctgctgttttttttaatacaaaagagggtttctcctcgatttcttcaataaaaaa 2700  
Db 2641 ctgctgttttttttaatacaaaagagggtttctcctcgatttcttcaataaaaaa 2700  
QY 2701 aaaaaa 2760  
Db 2701 aaaaaa 2760  
QY 2761 aaaaaa 2771  
Db 2761 aaaaaa 2771

RESULT 2  
US-09-508-377-11  
; Sequence 11, Application US/09508377  
; GENERAL INFORMATION:  
; APPLICANT: KALEEN, ZHONGYILI  
; APPLICANT: MORELL, MATTHEW  
; APPLICANT: RAHMAN, SADEOUR  
; TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS  
; FILE REFERENCE: 054270/0126  
; CURRENT APPLICATION NUMBER: US/09/508,377  
; CURRENT FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: AU PP 2509  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: PCT/AU98/00743  
; PRIOR FILING DATE: 1998-09-11  
; PRIOR APPLICATION NUMBER: AU PP 9108  
; PRIOR FILING DATE: 1997-09-12  
; NUMBER OF SEQ ID NOS: 71  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 2662  
; TYPE: DNA  
; ORGANISM: Triticum tauschli  
US-09-508-377-11

Query Match 91.7%; Score 2541.8; DB 8; Length 2662;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 2597; Conservative 0; Mismatches 32; Indels 6; Gaps 3;

QY 25 ctcccactttctctcccgccgacacccagtgctgacccggtctcatccaccatcacctcg 84  
Db 2 ctcccactttctctcccgccgacacccagtgctgacccggtctcatccaccatcacctcg 61  
QY 85 gctcggccacgggaaaccccccgatccgcttttggcagcgccactaaacccccggg 144  
Db 62 gctcggccacgggaaaccccccgatccgcttttggcagcgccactaaacccccggg 121  
QY 145 gagcgcccccggcgagcagcagcaccagtgaggagagagaggttgcgccccggcccg 204  
Db 122 gagcgcccccggcgagcagcagcaccagtgaggagagagaggttgcgccccggcccg 181  
QY 205 caccgagcgggcgatcccaccgtccgtgctcgacacctctccgctctccctgtccc 264  
Db 182 caccgagcgggcgatcccaccgtccgtgctcgacacctctccgctctccctgtccc 241

QY 265 cgcgccgcccacaccatggcgagcgagcgctcgagcgccgggtgctctcgccccagcgtc 324  
Db 242 cgcgccgcccacaccatggcgagcgagcgctcgagcgccgggtgctctcgccccagcgtc 301  
QY 325 cgctcgcgcccgatcccgcgacggcgcccgcgctcccgctctgctgcgtccgcgcgcgg 384  
Db 302 cgctcgcgcccgatcccgcgacggcgcccgcgctcccgctctgctgcgtccgcgcgcgg 361  
QY 385 ctccggcgcttggcgcgggcgcccgctacgtcgccgagctcagcaggaggagggcccgcgcg 444  
Db 362 ctccggcgcttggcgcgggcgcccgctacgttgcgagctcagcaggaggagggcccgcgcg 421  
QY 445 cgccccggcgagcagcagcaaatggcccccgctgctgcaggtctcccgcgcgcccg 504  
Db 422 cgccccggcgagcagcagcaaatggcccccgctgctgcaggtctcccgcgcgcccg 481  
QY 505 ccgccccggcccccgctcgccggcccccgacgcagcgccccctgcggagcgccggcggtg 564  
Db 482 ccgccccggcccccgctcgccggcccccgacgcagcgccccctgcggagcgccggcggtg 541  
QY 565 ggggaaactcgcccgacacctctctgcgaaggattgctgaggttccatcgacacacata 624  
Db 542 ggggaaactcgcccgacacctctctgcgaaggattgctgaggttccatcgacacacata 601  
QY 625 attgtggtgcaagtgcagcaggtattctgcagatcatgagtcgcaatgagcaacctcaagct 684  
Db 602 attgtggtgcaagtgcagcaggtattctgcagatcatgagtcgcaatgagcaacctcaagct 661  
QY 685 aaagtacacgtagcatcgctgttggactgggtgaagctgctctctatgcaaaagtcaagg 744  
Db 662 aaagtacacgtagcatcgctgttggactgggtgaagctgctctctatgcaaaagtcaagg 721  
QY 745 ggggtggagagatgttgggttgcgttaccgaattgctctgctgctgctgctgctgctgctg 804  
Db 722 ggggtggagagatgttgggttgcgttaccgaattgctctgctgctgctgctgctgctgctg 781  
QY 805 atggttgaatgccaagatacttaaatgggtctctctgataaaaaactatgcaaaaggcatt 864  
Db 782 atggttgaatgccaagatacttaaatgggtctctctgataaaaaactatgcaaaaggcatt 841  
QY 865 tacactggaagcacattaagattccattgcttggggggtacacatgaagtgcactttttt 924  
Db 842 tacactggaagcacattaagattccattgcttggggggtacacatgaagtgcactttttt 901  
QY 925 catgagatagagacaacgtcgatgtgggttgggtgcgatacatcgctcatatcacagacca 984  
Db 902 catgagatagagacaacgtcgatgtgggttgggtgcgatacatcgctcatatcacagacca 961  
QY 985 ggaagtattataggagataaattttgggtgcttttgggtgataatacagttcagatacacactc 1044  
Db 962 ggaagtattataggagataaattttgggtgcttttgggtgataatacagttcagatacacactc 1021  
QY 1045 ctttgcgtgctgcatgcgagggcccccaacttaactcttgaattgggaggtatatttatgga 1104  
Db 1022 ctttgcgtgctgcatgcgagggcccccaacttaactcttgaattgggaggtatatttatgga 1081  
QY 1105 cagaattgcattgttgcgtgacgattggcagcagccttgcgcagctcttgcctctcttgcct 1164  
Db 1082 cagaattgcattgttgcgtgacgattggcagcagccttgcgcagccttgcctctcttgcct 1141  
QY 1165 gcaaaatagaccatacgggttttacagagattcccgagacaccttgttatataataat 1224  
Db 1142 gcaaaatagaccatacgggttttacagagattcccgagacaccttgttatataataat 1201  
QY 1225 tttagcacatcaggggtgagagcctgcaagtacatactctgcgtggtggtggtggtggtggtg 1284  
Db 1202 tttagcacatcaggggtgagagcctgcaagtacatactctgcgtggtggtggtggtggtggtg 1261  
QY 1285 tggatgagcgttttagaattgggtatttccagaatgggagagagcagcagccttgcacag 1344  
Db 1262 tggatgagcgttttagaattgggtatttccagaatgggagagagcagcagccttgcacag 1321  
QY 1345 ggtgaggcaggttaacttttgaaggagcagttgtgacagcagatcggtattgtgaccggtc 1404







[illegible]







```
QY 1262 ctgatctgggattgctcctcctgaatgggtatgagcttttagatgggtattccagaatggg 1321
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 990 CTGACCTTGGTTGCCACCTGAATGGTATGGAGCTCTGGAGTGGGTATTCCTCGAATGGG 1049
QY 1322 caaggagcctgcctctgacaagggtgagcagttacactttttgaaaggagcagttgtga 1381
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1050 CGAGGAGGCTATGCCCTTGACAAGGGTGAGCAGTTAATTTTGAAGGTTGCAGTTGTGA 1109
QY 1382 cagcagatcgattgtgacagctcagtcaggtttattcatggggaggtcaacaactgctgaag 1441
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1110 CAGCAGATCGAATCGTGACTGTGAGTAAGGGTTATTATCATGGGAGGTCAACAATCTGCTGAAG 1169
QY 1442 gtggacagggcctcaatgagctcttaagctcccgcaaaaagtattgaatgggaattgtaa 1501
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1170 GTGCACAGGGCCCTCAATGACCTCTTAAGCTCCAGAAAGAGTGATTAAACGGAAATTTGTAA 1229
QY 1502 atggaattgacataatgattgaaacccaccacagacaagtgtctccctcatcattatt 1561
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1230 ATGGAATTGACATTAATGATTTGAACCCCTGCCACAGACAAATGTATCCCTGTCTATTATT 1289
QY 1562 ctgctgatgacctctctgaaagcccaaatgtaaaactgaattgcaagaaggagttgggtt 1621
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1290 CTGTTGATGACCTCTCTGNAAGGCCCAATGTAAAGGTGCAITTCAGAGAGGAGCTGGGTT 1349
QY 1622 tacctgtaaggaggatgttctctctgattggctttattggaagactggattaccagaaga 1681
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1350 TACCTATAAGGCCCTGATGTTCTCTCTGATTTGGCTTTATTGGAAGATTGGATTATCAGAAAG 1409
QY 1682 gcatatcatcctcaataatggccattccagagctcagagtcagtgaggagcgtgcgaattgtca 1741
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1410 GCATTGATCTATTCAACATTTATCATACCAGATCTCAATGCGGGAAGATGTTCAATTTGTCA 1469
QY 1742 tgcctggatctgggggtcccaatttttgaaggctggatgagatctaccgagtcgagttaca 1801
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1470 TGCTTGGATCTGTGTGACCCAGAGACTTGAAGATTGGATGGATGAGATCTACAGAGTCGATCTCA 1529
QY 1802 agataaattccgtgagtggttgattagtttccagtttccacacagaataaactgcag 1861
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1530 AGGATAAATTTCTGTGATGGGTGGATTTAGTGTCTCCAGTTTCCACCCACCAATAACTGCCG 1589
QY 1862 gttgcgatatattgttaatggcatcgagatttgaaccttgcgtcttcaatcagctatatg 1921
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1590 GCTGCCGATATATTGTTAATGGCATCCAGATTCCGAACCTTGTGGTCTCAATCAGCTATATG 1649
QY 1922 ctatgcaatatggtacagttctctgtattctatggaaactgggggctcccgagacacagtcg 1981
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1650 CTATGCAGTATGCCACAGTTCTCTGTTGTCATCCAACTGGGGGCTTAGAGATACCGTGG 1709
QY 1982 agacctcaaccttttgggtgcaaaaggagagagggtacaggggtgaggggtctctcacccg 2041
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1710 AGAACTTCAACCCCTTTCGGGTGAGAAATGGAGAGCAGGGGTACAGGGGTGGGCATTTCCACCCC 1769
QY 2042 taacctggtgacaagatggttggggcatttgcgaaccgcgcatgctcgacattcagggagcaca 2101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1770 TAACCACACAAACATGTTGTGCACATTCGGAACATGCAATATCTACATACAGGGGAACACA 1829
QY 2102 agccgtctctggaggggctcctaagcagggcatgacgaagaaccatacgtggggaccatg 2161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1830 AGTCTCTCTGGGAAGGGCTTAATGAAGCAGGAGCATGTCAAAAGACTTCACGTGGGACCATG 1889
QY 2162 cc-cgcagcagtcagcagcagatcttcagtggggcttcgtggaccaacctcagtcagt 2220
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1890 CCGCTGAACAAATACCAACAAATCTTCCAGTGGGCTTTCATCGATGACCCCTATGTCATGT 1949
QY 2221 agacgggggagtgagggt 2239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1950 AAAAAAGGACCAAAAGTGGT 1968
```

RESULT 5

US-09-508-377-13

; Sequence 13, Application US/09508377

```
; GENERAL INFORMATION:
; APPLICANT: KALEEN, ZHONGYILI
; APPLICANT: MORELL, MATTHEW
; APPLICANT: RAHMAN, SADEOUR
; TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS
; FILE REFERENCE: 054270/0126
; CURRENT APPLICATION NUMBER: US/09/508,377
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: AU PP 2509
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: PCT/AU98/00743
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: AU PP 9108
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 10337
; TYPE: DNA
; ORGANISM: Triticum tauschii
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (10232)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-09-508-377-13
```

```
Query Match 20.5%; Score 567.6; DB 8; Length 10337;
Best Local Similarity 93.7%; Pred. NO. 2.8e-72;
Matches 613; Conservative 0; Mismatches 39; Indels 2; Gaps 2;

QY 2056 atgttctgggcttcggaacgcgcatgtcgacattcaggagacacacccgttcctggag 2115
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9440 attgtcgaggcattgcgaacgcgcatgtcgacattcaggagacacacccgttcctggag 9499
QY 2116 gggctcatgaagcgggcatgacgaagaccatacgtggaccatgcc-cgcagcagtac 2174
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9500 gggctcatgaagcgggcatgacgaagaccatacgtggaccatgcccgccgagcagtac 9559
QY 2175 gagaagatcttcgagtgggccttcggaacacccctacgtatgacggggactggg 2234
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9560 gagcagatcttcgaatggccttcggaacacccctacgtatgacggggactggg 9619
QY 2235 gaggtccagtgagtcctctcctcagctctgaagacatcctctcctcctcccgcc 2294
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9620 gaggtcgaagcggggtctctctcctcagctctgaagacatcctcctcctcccgcc 9679
QY 2295 ggaaggataccctctacattgctgttcctgctacagtagagtcgaatgcgctgctt 2354
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9680 ggaaggataccctctacattgctgttcctgctacagtagagtcgaatgcgctgctt 9739
QY 2355 gctttggtcccggttcgagaacatacagcgctgtgctgctgctgctgctgctgctt 2414
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9740 gc-tgggtcccggttcgagagtagatgacggtgtgctgctgctgctgctgctgctt 9798
QY 2415 ggggtgagcaggttacagttttggggaataaggaagagatgctgctgagatggttaac 2474
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9799 ggggtgagcaggttacagttttggggaataaggaagagatgctgctgagatggttaac 9858
QY 2475 agcaaaagcaccactcagatggcagctctctctgctgctgctgctgctgctgctgctt 2534
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9859 agcaaaagcaccactcagatggcagctctctctgctgctgctgctgctgctgctgctt 9918
QY 2535 actggtgactcttagcttagtgatgagtttggcttctgcttctgctgctgctgctt 2594
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9919 actggtgactcttagcttagtgatgagtttggcttctgcttctgctgctgctgctt 9978
QY 2595 gtccttagtgacaaatatttgacctgttggaagaatttatcttctgctgctgcttctt 2654
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9979 gtccttagtgacaaatatttgacctgttggaagaatttatcttctgctgctgcttctt 10038
QY 2655 taatcaaaagagggggttctcctccgatttcaataaaaaaataaaaaa 2708
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```











Db 1 catcggggaagatgttcaatttgtcattgtctggtatctggtgacccagaagcttgaagagtg 60

Qy 1776 gatgagatctaccgagtcgagttacaaggataaattccgtggatgggttggaattagtgt 1835

Db 61 gatgagatctacagtcgagttctcaaggataaattcgtggatgggttggaattagtgt 120

Qy 1836 tccagtttccacagaataactcgagtttcgatatattgttaatgcgcattcgagatttga 1895

Db 121 tccagtttccacacggaataactgcgggttcgatatattgttaatgcgcattcgagatttga 180

Qy 1896 accttggggtcttaatcagctatatgtctatgcataatgatgttcagtttcctgtagttcatcgg 1955

Db 181 accttgtgtctcaatcagctatatgtctatgcagtatggcagttctctgtgtctcatcgc 240

Qy 1956 aacttgggggtcttcgagacacagtcgagacttcaacccttttgggtcgaaaggagagga 2015

Db 241 aacttgggggtcttcgagataccgttggaagaacttcaaccctttcgggtgagaatgggagagca 300

Qy 2016 ggggtacaggggtggggtctctcacgcctaac--gtggacaagaatgtgttgggcatttgcgaa 2074

Db 301 ggggtacaggggtggggtctctcacgcctaacacagaaaacatgttgttggacatttgcgaa 360

Qy 2075 ccgcgatgtcacatttcagggagcacaagccgtctctggagggggtcctatga 2125

Db 361 ctgcaatctctacatcaggaagaacagactctctctgagaagagctaatga 411

```

RESULT 11
US-10-021-323-969
; Sequence 969, Application US/10021323
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 969
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3825-016-Q1-K6-D4
US-10-021-323-969

```

	Query Match	11.1%	Score	307.6;	DB	9;	Length	598;
	Best Local Similarity	71.7%	Prod. NO.	3.3e-35;				
	Matches	403;	Conservative	0;	Mismatches	159;	Indels	0;
QY	1448	agggcctcaatgagctctaaagctcccgaaaagtgattgaaagggaattgtaaatgaa	1507					
Db	12	atgggttaaatcagctattaaacacgctcgaggagtggtctgactggaattacaaacggca	71					
QY	1508	ttgacattaatgattggaaacccacacagacaagtgctccctcatcatattctgtctg	1567					
Db	72	ttcatattgctgaatgggactccattccagatgagcatatgcttcccatcatctgctg	131					
QY	1568	atgacctctctggaaaggcccaatgtaaagctggaattgcagaaaggagtggtggttacctg	1627					
Db	132	cttctctatctggaaaggttcaatgcagagactgctctccacagggaacttgggtttccaa	191					
QY	1628	taaggagagatgtctctctgattgctcttatttggaaagactggattaccagaaagccattg	1687					
Db	192	taaggctgattgctctctgattgatttcatcggaagactggagactcaccagaaagccattg	251					
QY	1688	atctcattaaaaatggccattccagagctcatgaggaggagactgcaatttgtcagctgtg	1747					

[illegible]

RESULT 12  
US-09-508-377-15  
; Sequence 15, Application US/09508377  
; GENERAL INFORMATION:  
; APPLICANT: KALEEN, ZHONGYILI  
; APPLICANT: MORELL, MATTHEW  
; APPLICANT: RAHMAN, SADEOUR  
; TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS  
; FILE REFERENCE: 054270/0126  
; CURRENT APPLICATION NUMBER: US/09/508,377  
; CURRENT FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: AU PP 2509  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: PCT/AU98/00743  
; PRIOR FILING DATE: 1998-09-11  
; PRIOR APPLICATION NUMBER: AU PP 9108  
; PRIOR FILING DATE: 1997-09-12  
; NUMBER OF SEQ ID NOS: 71  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 5072  
; TYPE: DNA  
; ORGANISM: Triticum tauschii  
US-09-508-377-15

	Query Match	10.8%;	Score 299;	DB 8;	Length 5072;
	Best Local Similarity	100.0%;	Pred. No. 7e-34;		
	Matches 299;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	cgcacctcacctgcgcttgccacactccacactctctctctcccgcgcacacccagatcgcc	60		
Db	4774	cgcacctcacctgcgcttgccacactccacactctctctctcccgcgcacacccagatcgcc	4833		
QY	61	accggtctatcccatcacctctggcctcgccacccggaacccccccgatacgccttttg	120		
Db	4834	accggtctatcccatcacctctggcctcgccacccggaacccccccgatacgccttttg	4893		
QY	121	caggcgcgcactaaacccccggggagcgcgcccccgcgacgacgacccgcagtggg	180		
Db	4894	caggcgcgcactaaacccccggggagcgcgcccccgcgacgacgacccgcagtggg	4953		
QY	181	agagagagcttcgccccggcccgacccgagcgggcgatccacacgtcgtccgcga	240		
Db	4954	agagagagcttcgccccggcccgacccgagcgggcgatccacacgtcgtccgcga	5013		
QY	241	ctctctcgcctctctccctgtcccgcgccacacccatggcgcgacggcgctgg	299		
Db	5014	ctctctcgcctctctccctgtcccgcgccacacccatggcgcgacggcgctgg	5072		



```

; PRIOR APPLICATION NUMBER: US 60/161,619
; PRIOR FILING DATE: 1999-10-26
; NUMBER OF SEQ ID NOS: 17472
; SEQ ID NO 17004
; LENGTH: 635
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3279-2
US-09-696-664A-17004

```

Query Match 10.6%; Score 293.6; DB 7; Length 635;  
Best Local Similarity 85.0%; Pred. No. 3.3e-33;  
Matches 340; Conservative 0; Mismatches 59; Indels 1

```

RESULT 15
US-09-849-529A-17365
; Sequence 17365, Application US/09849529A
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51893)B
; CURRENT APPLICATION NUMBER: US/09/849,529A
; CURRENT FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: US 60/196,868
; PRIOR FILING DATE: 2000-05-09
; NUMBER OF SEQ ID NOS: 24076
; SEQ ID NO 17365
; LENGTH: 577
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3584-019-Pl-K1-G3
US-09-849-529A-17365

```

```

RESULT 13
US-09-865-439A-21857
; Sequence 21857, Application US/09865439A
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51936)B
; CURRENT APPLICATION NUMBER: US/09/865,439A
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,458
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 119126
; SEQ ID NO 21857
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3353-011-Pl-K1-C1
US-09-865-439A-21857

```

Query Match	10.6%	Score 293.6	DB 7	Length 412
Best Local Similarity	85.2%	Pred. No. 3.2e-33		
Matches 351	Conservative 0	Mismatches 59	Indels 2	Gaps 2
QY 1203	cagcacccttggtatatacataattgagcacatcaaggggtgtggagcctgcaagatcacatccc	1262		
Db				
	1 cagcattctgtgaatacataattgagcacatcttgtgtatagcctgcaagcacatctgc	60		
QY 1263	tgatc-tgggaattgccctctcctgaaatggatggagctttlagaatgggtattttccagaatggg	1321		
Db				
	61 tgggcgtgtgggttgcaccctgtatggtatgtaacttggagtggtattccctgaatggg	120		
QY 1322	caaggagggcatccctcttgacaaaggtgaggcagtgtaactttttgaaaggagcattgttga	1381		
Db				
	121 cgaggagggcatgccctctgacaaaggtgaggcagtgtaatttttgaagggtgcattgttga	180		
QY 1382	cagcagatcggattgttgaccctgcagtcaggggtattcatctggaggtcacaaactgctgaag	1441		
Db				
	181 cagcagatcgaatcgtgactgtcagtaagggtattcttgggtggtcacaaactgctgatg	240		
QY 1442	gtgacagggccctcaatgagctctttaaagctcccgaaaaagtgtattgaaatggaattgttaa	1501		
Db				
	241 ggggacagggccctcaatgagctctttaaagcttcagaaagggtatttaaagggaattgtaa	300		
QY 1502	atggaattgacattaaatgattgaaacccaccacagacaaagtgcttcccctcatcata-t	1560		
Db				
	301 atggaattgacattatgattgaaacccctgcagacgaatgtatactgtctcattatt	360		
QY 1561	tcgtcgatgacctctctgaaaggccaaatgttaaagctgaaatgcagaagg	1612		
Db				
	361 tcctgttgatgacctttttgaaaggccaaatgttaaaggtgcattgtcgagaagg	412		

```

RESULT 14
US-09-696-664A-17004
; Sequence 17004, Application US/09696664A
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Andersen, Scott E.
; APPLICANT: Dubois, Patrice
; APPLICANT: Mahadeo, Debbie A.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: plants
; FILE REFERENCE: 38-21(51721)B
; CURRENT APPLICATION NUMBER: US/09/696,664A
; CURRENT FILING DATE: 2000-10-25

```



```

Db      1 ccaaataccgtccatttgggtttataaagatgctcggagcgcccttgtaattcataacc 60
Qy 1226 tagcacatcagggtgtggaagccttgcaagtacatatcctgatctgggagtcctccctgaat 1285
Db      61 ttgcgatcaggagtggaaccttgagcaaccttttaagaaatttgggattgcccctcaact 120
Qy 1286 ggtatggagcttttagaatgggtatttccagaaatgggcaaggagggcatgccccttgacaaag 1345
Db     121 ggtatggggcattggaatgggtgttccctacatgggcaaggacacatgctcttgacacag 180
Qy 1346 gtgaggcagtttaactttttgaaaggagcagttgtgacagcagatcggattgtgaccgtca 1405
Db     181 gagaggctgtcaattattctaaagggtgccattgtgacatctgacatctgcaatactaccgtta 240
Qy 1406 gtcagggttattcatgaggaggtcacaaactctgaagtgacagggcctcaatgagctct 1465
Db     241 gtaagggtctatgcatggaatggaatacaaaactgttgaaggcgatggttttaaatcagctat 300
Qy 1466 taagctcccgaataaagtgtattgaatggaattgtaaatggaattgacattaatgattgga 1525
Db     301 taagcagtcggaggagtgcttgactggaattacaaacggcattcatattgctggaatggg 360
Qy 1526 accccacacagacaagtgtctccctcatcattattctgtcgatgacctctctggaaaagg 1585
Db     361 atccatctcagatgagcatattgcttccattattctgtctgtctctctctctctctctct 420
Qy 1586 ccaaatgaaagctgaattgcagaagagttgggtttaccctgtaaggaggaggtgtctctc 1645
Db     421 ttcattgcaagactgctctccacaggaacttgggtttccaattaaagcctgggtgtctcct 480
Qy 1646 tgattggctttattggaagactggtattacc-agaaggcattgattcatctcaataaaggcc 1704
Db     481 tgatgggatttattcggagactggaactccaagaaggcattgaccttgatcactgggca 540
Qy 1705 attccagagctcatgaggaggagcgtgcaatttctca 1741
Db     541 attccagaacttatggaagatgactacaaaattgtaa 577

```

Search completed: March 28, 2002, 15:15:52  
Job time: 7581 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2002, 14:55:35 ; Search time 71.31 Seconds  
(without alignments)  
785.295 Million cell updates/sec

Title: US-09-674-824-2  
Perfect score: 4044  
Sequence: 1 MAATGVCAGCLAPSRLRAD.....SDGSLSVRVTAETRNQLVTL 756

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues  
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	4044	100.0	756	21	AA150818	Wheat soluble star
2	3612	89.3	671	19	AAW23937	Wheat soluble star
3	3348	82.8	647	20	AA109004	Wheat starch solub
4	2749	68.0	649	19	AAW38218	Maize starch synth
5	2609.5	64.5	626	15	AA151231	Soluble rice starch
6	2557.5	63.2	583	19	AAW56488	Zea mays soluble s
7	2548.5	63.0	583	19	AAW70894	Maize starch solub
8	2390.5	59.1	539	19	AAW58491	Zea mays starch sy
9	1718.5	42.5	459	17	AA199540	Soluble starch syn
10	1205.5	29.8	698	19	AAW56487	Zea mays soluble s
11	1203.5	29.8	804	19	AAW70892	Maize starch solub

12	1174.5	29.0	799	21	AA137567	Wheat starch synth
13	1174.5	29.0	799	21	AA137597	Wheat starch synth
14	1165	28.8	798	21	AA137566	Wheat starch synth
15	1148.5	28.4	799	19	AAW23938	Wheat granule-boun
16	1120.5	27.7	597	21	AA137568	Wheat starch synth
17	1068	26.4	669	19	AAW70893	Maize starch solub
18	1068	26.4	669	19	AAW56486	Zea mays soluble s
19	877	21.7	609	19	AAW56485	Oryza sativa starch
20	876	21.7	606	13	AA125476	Granule-bound star
21	869	21.5	502	13	AA125474	Central fragment f
22	865	21.4	533	19	AAW56484	Zea mays waxy gene
23	865	21.4	637	19	AA116604	Protein encoded by
24	863	21.3	563	21	AA104667	Arabidopsis thalia
25	863	21.3	610	21	AA104666	Arabidopsis thalia
26	862.5	21.3	527	21	AA104668	Arabidopsis thalia
27	769.5	19.0	209	21	AA113555	Arabidopsis thalia
28	714.5	17.7	193	21	AA113556	Arabidopsis thalia
29	678.5	16.8	477	13	AA125462	glgA, Escherichia
30	678.5	16.8	477	19	AAW53890	E. coli glycogen s
31	632.5	15.6	495	19	AAW70885	Protein encoded by
32	617.5	15.3	277	19	AAW5849	S. pneumoniae deri
33	556	13.7	677	17	AA199539	Soluble starch syn
34	556	13.7	1230	18	AAW17785	Potato tuber solub
35	556	13.7	1230	21	AA149306	Potato starch synt
36	551	13.6	735	18	AAW27229	Z. mays starch syn
37	511.5	12.6	534	21	AA149307	Wheat starch synth
38	498.5	12.3	1059	21	AA137570	Wheat starch synth
39	495.5	12.3	1828	21	AA137569	Wheat starch synth
40	495.5	12.3	1628	21	AA149304	Wheat starch synth
41	492	12.2	1674	20	AA106199	Maize endosperm st
42	492	12.2	1674	21	AA149305	Maize starch synth
43	469.5	11.6	265	21	AA141354	Arabidopsis thalia
44	448.5	11.1	1170	22	AA170781	Maize starch synth
45	431.5	10.7	293	22	AAU00036	Wheat Starch synth

ALIGNMENTS

RESULT 1  
AA150818  
ID AA150818 standard; Protein; 756 AA.  
XX  
AC AA150818;  
XX  
DT 18-FEB-2000 (first entry)  
XX  
DE Wheat soluble starch synthase protein.  
XX  
KW Soluble; starch synthase; wheat; transgenic plant; starch production;  
KW food; baking; pastry; packaging material; glucose; glucan; paper; pulp;  
KW adhesive; textile; building material; soil stabilizer; wetting agent;  
KW fertilizer; plant-protection; cosmetic; flocculant.  
OS Triticum aestivum.  
XX  
PN DE19820607-AL.  
XX  
PD 11-NOV-1999.  
XX  
PF 08-MAY-1998; 98DE-1020607.  
XX  
PR 08-MAY-1998; 98DE-1020607.  
XX  
PA (AGRE ) HOECHST-SCHERING AGREVO GMBH.  
XX  
PI Loerz H, Luetticke S, Block M;  
XX  
DR WPI; 2000-024508/03.  
XX  
N-PSDB; AA224487.  
XX  
PT New enzyme with starch synthase activity, useful for producing starch  
for foods and packaging materials -



XX Claim 1a: Page 19-21; 24pp; German.  
 XX This invention describes a novel protein (I) with the activity of wheat  
 CC starch synthase. Transgenic plants, specifically wheat, that contain (I)  
 CC are used for production of starch, used particularly in foods,  
 CC particularly baked and pastry goods and for making packaging materials or  
 CC disposable items. Starch may also be used as starting materials for  
 CC glucose or glucan components (e.g. for fermentation or further chemical  
 CC conversion); in paper and pulp production, as adhesives, in textiles,  
 CC in preparation of gypsum-based building materials, as soil stabilizer,  
 CC as wetting agent etc. in fertilizer and plant-protection compositions,  
 CC as binder (in pharmaceuticals, cosmetics, coal briquetting and casting  
 CC sand), as flocculant in soil or coal slurries, as rubber and leather  
 CC additives, and for production of synthetic polymers, e.g. polyurethane  
 CC films. Transgenic plants with increased/decreased production of (I)  
 CC produce starches with altered physical and/or chemical properties such as  
 CC amylose/amylopectin ratios, degree of branching, mean chain length,  
 CC phosphate content, gelatinization properties, gel- or film-forming  
 CC properties, or starch grain size or structure. This sequence represents  
 CC the soluble starch synthase isolated from wheat (Triticum aestivum L. cv.  
 CC Florida).  
 XX  
 XX Sequence 756 AA;

Query Match 100.0%; Score 4044; DB 21; Length 756;  
 Best Local Similarity 100.0%; Pred. NO. 0;  
 Matches 756; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAATGVCAGCLAPSVRLRADPATAARASACVVRARLRRLARGRYVAELSGEGPAARPAQ 60  
 Db 1 maatgvagclapsvrlraopacaarasacvvrarrrlarygrvaelsgregpaarpag 60  
 QY 61 QQLAPPLVPGFLAPPPAPAPQSPAPTQPLPDAGVAGELAPDLLLEGIAEDSIDSIIVAAAS 120  
 Db 61 qqlapplvpgflappppapapqspaptqplpdagvagelapdlllegiaedsidsiivaas 120  
 QY 121 EQDSEIMDANEQPAQAKVRSIVFVTGEAPYAKSGGLGVCGSLPTALAAARGHRVWVMP 180  
 Db 121 eqdseimdanepqakvrsivfvtgeapayaksgglgvcgsplalaaarghrvwnvmp 180  
 QY 181 RYLGSSDKNYAKALYAKHAKIKPCFGSGSHEVTFEYRDNDVWFVDHPSYHRPGSLYG 240  
 Db 181 rylngssdknyakalyakhakikpcfgsgshevtfeyrdndvfwfvdhpsyhrpgslyg 240  
 QY 241 DNFAGCDNQFRTLCLCYAACEAPLLELGGYIYGQNCMFVNDWHASLVVPLLAARYP 300  
 Db 241 dnfgagcdnqfrytlclcyaaaceaplillelgyiygqncmfvndwhaslvpvllaakyrp 300  
 QY 301 YGYVRDSRLTVLTHNAHOGVEPASTYDGLPPEYKGALEWVPEWARHDLKGEAVN 360  
 Db 301 ygyvrdsrslvtlthnahogvepastydgldppeykgalewvpewarhaldkgeavn 360  
 QY 361 FLKGAVVTAADRIVTSQGYSWEVTTAEGGOGNELLSSRKSVLNGIVNGIDINDWNPSTD 420  
 Db 361 flkgavvtadrlvtvsqgyswvtttaeggggnellssrksvlingvngidindwnpttd 420  
 QY 421 KCLPHYSVDDLSGKAKRAELQELGLPVREDVPLIGTIGRLDYQKGLIDIKMAIPELM 480  
 Db 421 kclphysvddlsgakkaelqelglpvredvpiligtigrldyqkgldikmaiipelm 480  
 QY 481 REDVQFVMLGSGDPIFEGNMRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFEP 540  
 Db 481 redvqfvmllsgdplifegnmrstessykdfrgwgfsvpvshritagcdillmpsrfe 540  
 QY 541 CGLNQLYAMQYGRVPVYVHGTGGLRDVTFEFPFGAGEGCTGWAFLSPLTVDKMLWALRTA 600  
 Db 541 cglnqlyamqygrvpvyvhtgtgglrdvtfefpfgagegctgwafslpltdkmlwalrta 600  
 QY 601 MSTFRHKKPSWGLMKRGMTKHTWDHAPSSRSRSGSPSWTNPSTCRGLGRSKCESPS 660  
 Db 601 matfrehkpswglmkrgmtkhtwdhapsestsrsgspswtntpsctrrglgrskcesps 660

QY 661 ALKTSSSSFRGPEGYPCITLRCPATVESQCACLLWFAGSRTYDGCAAAATASGGRQLQFW 720  
 Db 661 alktsssfgrgpegyptcltrcpatvesqcacllwfagsrtydgcaaaavtasgrqlqfw 720  
 QY 721 GIRKGCAGWLTAKHHSDGSLSVRVTAETIRNQLVTL 756  
 Db 721 girkgcaagwltakhhsgslsvrvtaetirnlvltl 756

RESULT 2  
 AAW23937  
 ID AAW23937 standard; Protein: 671 AA.  
 AC AAW23937;  
 XX 21-MAY-1998 (first entry)  
 DT Wheat soluble starch synthase partial sequence.  
 DE Starch synthase; wheat; transgenic plant.  
 XX  
 KW Triticum aestivum L. cv. Florida.  
 OS  
 XX WO9745545-A1.  
 PN  
 XX 04-DEC-1997.  
 PD  
 XX 28-MAY-1997; 97WO-EP02793.  
 PF  
 XX 11-SEP-1996; 96DE-1036917.  
 PR  
 XX 29-MAY-1996; 96DE-1021588.  
 PR  
 XX (AGRE ) HOECHST-SCHERING AGREVO GMBH.  
 PA  
 XX Block M, Loerz H, Luetticke S, Froberg C, Kossmann J;  
 PI Walter L;  
 PI  
 XX WPI; 1998-032652/03.  
 DR N-PSDB; AAV01527.  
 DR  
 XX Nucleic acid encoding starch synthase enzymes from wheat - for  
 PT transgenic plants that produce modified forms of starch, useful e.g.  
 PT in foods, or for production of packaging materials and disposable  
 PT goods

Claim 1; Page 47-50; 71pp; English.  
 This amino acid sequence comprises a near full-length sequence for  
 a soluble starch synthase of summer wheat (cv. Florida). It was  
 deduced from a cDNA clone (AAV01527) isolated from a 21-day caryopsis  
 cDNA library. A granule-bound starch synthase (see AAW23938) has  
 also been identified. Isolated nucleic acids encoding these  
 enzymes can be inserted into vectors for production of transgenic  
 plants, particularly starch-producing plants, specifically wheat.  
 Use of the isolated nucleic acids, or of antisense sequences, allows  
 starch metabolism to be regulated in transgenic plants.  
 Overexpression may result in improved crop yield, while modification  
 of starch in planta may eliminate the need for subsequent  
 chemical/physical modification. Plants with altered levels of the  
 various isoforms of starch synthase will produce starch of different  
 chain length, amylose/amylopectin ratio, degree of branching,  
 phosphate content, gelatinisation behaviour, granule size and shape,  
 viscosity etc. The starch produced by such plants is useful  
 particularly in foods (especially bakery goods or pasta) or to  
 produce packaging materials or disposable goods, as well as in any  
 other known use of starch.  
 XX Sequence 671 AA;

Query Match 89.3%; Score 3612; DB 19; Length 671;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-301;



Matches 671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	86	TOPPLPDAGVGLAPDLLLEGIAEDSIDSIIVAASQDSEIMDANEOPOAKVTRTSVEVTT	145
Db	1	tcppldagvgelapdllleglaedsidsiivaaseqdeimdanecpqaakvtrtsivft	60
QY	146	GEAAPYAKSGGLDVCGLSPALAAARGHRVWVMPRYLNGSSDKNYAKALYTAKHKIFPC	205
Db	61	geaapyaksgglgdvcgslpialaarghrvmvmprylmgssdknyakalytakhkikpc	120
QY	206	FGGSHVEVTFEHEYRDNVDFVFDHPSYHRPGSLYGNDFGAFGDNQRPYTLCCYAACEAPL	265
Db	121	fggshvevtfheyrdndvdfvfdhpsyrpgslygndfgafgdnqfrytlccyaaceapl	180
QY	266	ILELGGYIYGQCMFVNDWHASLVPVLLAAKYRPYGVYRDSRSTLVIRNLAHOGVEPAS	325
Db	181	ilelgsyiygqcmfvndwhaslvppvllaakyrpygvyrdsrstlvirnlahogvepas	240
QY	326	TYPDGLPPEWYGALWVFPWEARRHALDKGEAVNFKGAVVTADRVTVSQGSWEVTT	385
Db	241	typdglppewygalewfpewarrhaldkgeavnflkgavvtadrvtvsqgswevtt	300
QY	386	AREGGQLNELLSRKS SVLNGIYNGIDINWNP TTKCLPHHSVDL SGKAKCKAE LQKE	445
Db	301	aegggqlnelllsrksvlnvgidindwmp ttdkclphhsvddls gkakckaelqke	360
QY	446	LGLPVREDVPLIGFIRLDYOKGIDLKMAIPELMREDYQVYMLGSGDPTFEQWMBSTES	505
Db	361	lglpvredvpllfgfirlgyqgldikmaipe lmedvqfvm lgsqdp lfeqwmrstes	420
QY	506	SYKDFRGVWGFSVPVSHRITAGCDILLMPSRFEPGCLNQLYAMQYGTVPVHGTGGLRD	565
Db	421	sykdkfrgvwgfsvpvshritagcdillm psrfepcglnqlyamqygtvpvhgtgglrd	480
QY	566	TYVETNPFGAKGEGTGWAFSP LTVDKMLWALRTAMSTPREHKPSWEGMLKRGMTKDHFW	625
Db	481	tvetnfpfakgeegtgwafsp ltvdkmlwalrtamstprehkp sweglmkrgmtkdhfw	540
QY	626	DIAPSTSRSSGSPSWNTPTSCRRGLGRSKCSPSALKTSSSFRGPEGYPCPTLRCPATV	685
Db	541	dhapestsrssgspswntptscrrglgrskcsp salktsssf rgepypcptlrcpatv	600
QY	686	ESQACLLWFAGSRTYDGC AAAAVTASGGRLQF WGIKRGCAAGWL TAKHHSDGSLSVRV	745
Db	601	esqaccllw fagsrtydgc aaavtasgrrlqf wgi rkgcaagwltak hsdgslsvrv	660
QY	746	TAEIRNQLVTL 756	
Db	661	tairnqlvltl 671	
RESULT 3			
AAY09004			
ID	AAY09004 standard; Protein; 647 AA.		
XX			
AC	AAY09004;		
DT			
XX			
XX	05-JUL-1999 (first entry)		
DE	Wheat starch soluble synthase I (SSS I) amino acid sequence.		
XX			
KW	Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBSS;		
KW	starch branching enzyme; starch soluble synthase; debranching enzyme;		
KW	endosperm; wheat; barley; granule-bound synthase; glutenin; starch;		
KW	grain softness protein I; bacterial isoamylase; glycogen synthase;		
XX	WSBE I-D4 gene.		
OS	Triticum tauschii.		
XX			
PN	W09914314-Al.		
XX			
PD	25-MAR-1999.		
XX			

PF	11-SEP-1998;	98WO-AU00743.	
XX			
PR	20-MAR-1998;	98AU-0002509.	
PR	12-SEP-1997;	97AU-0009108.	
XX			
PA	(CSIR ) COMMONWEALTH SCI & IND RES ORG.		
PA	(GOOD-) GOODMAN FIELDER LTD.		
PA	(LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.		
PA	(AUSU ) UNIV AUSTRALIAN NAT.		
XX			
PI	Li Z, Morell M, Rahman S;		
XX			
DR	WPI; 1999-229525/19.		
DR	N-PSDB; AAX34651, AAX34652.		
XX			
PT	New isolated cereal plant enzyme genes used for, e.g. expression of		
PT	antisense sequences of granule bound synthase		
XX			
PS	Claim 13; Page 95-97; 171pp; English.		
XX			
CC	The invention relates to a novel enzyme of starch biosynthetic pathway		
CC	in a cereal plant, where the enzyme is selected from starch branching		
CC	enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching		
CC	enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or		
CC	SBE I of rice or maize. The methods and products can be used for		
CC	targeting expression specifically to the endosperm of the seeds of cereal		
CC	plants such as wheat or barley. They can be used for the expression of		
CC	e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low		
CC	mol. wt. glutenin, grain softness protein I, bacterial isoamylase,		
CC	bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They		
CC	can be used for modifying the characteristics of starch produced by a		
CC	plant. The present sequence represents the wheat SSS I protein sequence.		
XX			
SQ	Sequence 647 AA;		
Query Match 82.8%; Score 3348; DB 20; Length 647;			
Best Local Similarity 99.7%; Pred. No. 1.9e-278;			
Matches 626; Conservative 1; Mismatches 1; Indels 0; Gaps 0;			
QY	1	MAATGVAGCLAPSVRLRADPATAARASACVVRARLRRLARGRYVAELSRGGAARPAQ 60	
Db	1	maatgvagclapsvrlradpataarasacvvrarlr largryvaelsregpaarpaq 60	
QY	61	QQLAPPLVCGFLAPPAPAPQSPAPTQPLPDAGVGLAPDLLLEGIAEDSIDSIIVAAS 120	
Db	61	qqlapplvpgflappapapsaptqpplpdagvgelapdlllegiaedsidsilvaas 120	
QY	121	EQDSEIMDANEQPAQVTRTSIVFTGAEAPYAKSGGLGVDYCGSIPALAAARGHRVWVMP 180	
Db	121	eqdseimdanecpqaqvtrtsivftgeaapyaks gglgvdvcgsalpialaarghrvmvmp 180	
QY	181	RYLNGSSDKNYAKALYTAKHIKIPCFGGSHVEVTFEHEYRDNVDFVFDHPSYHRPGSLY 240	
Db	181	rylngssdknyakalytgkhikipcfggshvevtf fheyrdndvdfvfdhpsyrpgsly 240	
QY	241	DNFGAFGDNQFRYTLCCYACEAPLILELGGYIYGQCMFVNDWHASLVPVLLAAKYRP 300	
Db	241	dnfgafgdnqfrytlccyaaceaplillelgsyiy gqcmfvndwhaslvppvllaakyrp 300	
QY	301	YGYVRDSRSTLVIRNLAHOGVEPASTYPDLGLPPEWYGALWVFPWEARRHALDKGEAVN 360	
Db	301	ygyvrdsrstlvirnlahogvepasty pdlglppew ygalwfpewarrhaldkgeavn 360	
QY	361	FLKGAVVTADRVTVSQGSWEVTTAEGGQGLNELLSRKS SVLNGIYNGIDINWNP TTD 420	
Db	361	flkgavvtadrvtvsqgswevtt aeggqglne llssrksvlnvgidindw np ttd 420	
QY	421	KCLPHHSYVDL SGKAKCKAE LQKE LGLPVREDVPLIGFIRLDYOKGIDLKMAIPELM 480	
Db	421	kc lphhsyvdls gkakckaelqke lglpvredv pligfirl dyokgidlkmaipelm 480	
QY	481	REDVQFVMLGSGDPIFEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFEP 540	



```
Db 481 redvqfvmlysggdpifegwmrtesykdfrgwgfsvpvshritagcdillmpsfef 540
QY 541 CGLNQLYAMQYGVVPHVHGTGGLRDTVETENPFCAGKEEGTGNFAFSLPTVDKMLWALRTA 600
Db 541 cglinqlyamqygtvpvvhggtggrldtvetfnpgakgeegtgwafagpltdvdkmlwalrta 600
QY 601 MSTFREHKPSWEGLMKRGMTKDHWDHA 628
Db 601 mstfrehkpsweglmkrmtkdhtwdha 628

RESULT 4
AAW38218
ID AAW38218 standard; Protein: 649 AA.
AC AAW38218;
DT 22-MAY-1998 (first entry)
XX Maize starch synthase type I.
DE Maize starch synthase type I.
XX Maize; starch synthase type I; starch.
XX Zea mays.
XX
XX PN DEL9619918-A1.
XX XX
XX PD 20-NOV-1997.
XX PF 17-MAY-1996; 96DE-1019918.
XX PR 17-MAY-1996; 96DE-1019918.
XX XX
XX PA (PLAN-) PLANTTEC BIOTECHNOLOGIE GMBH.
XX
XX PI Froberg C, Kossmann J;
XX XX
XX DR WPI: 1998-000821/01.
XX DR N-PSDB; AAT95785.
XX XX
XX PT DNA encoding maize starch synthase type I protein - for producing
XX transgenic plants.
XX PS Claim 1: Pages 16-20; 23pp; German.
XX
XX CC The present sequence is maize starch synthase type I, useful
XX in the production of starch. Starch can be used in various
XX conventional starch applications, e.g. starch hydrolysate products,
XX foods, papermaking, adhesives, textiles, building materials, soil
XX stabilisation, agrochemicals, pharmaceuticals, cosmetics, coal
XX briquettes, ore and coal slurries, foundry casting, rubber, leather
XX and synthetic polymers. The enzyme produces a starch stated to have
XX different physicochemical properties, especially viscosity and
XX gelling properties, from wild type starch.
XX
XX SQ Sequence 649 AA;

Query Match 68.0%; Score 2749; DB 19; Length 649;
Best Local Similarity 82.5%; Pred. No 5e-227;
Matches 518; Conservative 34; Mismatches 62; Indels 14; Gaps 4;

QY 4 TGVGAGCLAPSRLRADPATAARASACVVR-----RLRLRARGRYVAELSRGPAARPAQ 60
Db : ||| || : || : || : || : || : || : || : || : || : || : || : ||
14 savgaacll-----laraawaavagdrprlrlqrvlrrrcvaelsregpaprmp 65
QY 61 QQLAPPLVPGFLAPPAPPAQSPAPTOPPLPDAGVCELAPDLLEGIAEDSIDSIIVAA 120
Db ||||| ||| || : || : || : || : || : || : || : || : || : || : ||
66 allappilvpgfla--paaeptgepaltpppvpdaglvgve--peglaegsidntvvva 122
QY 121 EODSEIMDANEQPAQKAVTSIVVTGEAAPYAKSGGLGVCGSLPIALAAARGHRVWVMP 180
||| ||| : || : || : || : || : || : || : || : || : || : || : ||
```

```
Db 123 eqdseivvgkeqarakvtqnivfvtgeaspyaksgglgdcgslpvalaarghrvmvmp 182
QY 181 RYLNSSDKNAYKALYTAKHKIPCFGGSHVETFFEHYRDNVDWVFVDHPSYHRPGSLYG 240
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
Db 183 rylngtsdknyanafytekhiripcfgehevtffheyrdsvdwfvvdhpsyrpgnlyg 242
QY 241 DNFGAFGDNQFRYTLICYAACEAPLILELGGYIYGQCMFVVDWHAASLVPVLLAAKYP 300
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
Db 243 dkfgafgdnqfrytllyaaaceaplvlglggylygqcmfvvndwahaslvpillaakyrp 302
QY 301 YGVYRDSRSTLVITHNLAHQGVPEASTVPDLGLPPEWYGALEWVPEWARHDLKGEAVN 360
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
Db 303 ygvykdsrslvlhmlangqvepastypdlglpewygailewvfpewarhaldkgeavn 362
QY 361 FLKGAVVTADRIVTVSQYSWEVTTAREGGGLNELLSRSKSVLNGIYNGIDINDNPTTD 420
Db 363 flkgavvtadrivtvskgyswevttaeggglnellssrksvlnvngidindwnpatd 422
QY 421 KCLPHYSVDDLSGKAKKAELOKELGLPVREDYPLIGFGRDLQYKGLIDLKMAIFELM 480
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
Db 423 kclpchysvddlsgkakckgalqkelgiprdvpvlfigridyqkgldliqilpdlm 482
QY 481 REDVQFVMLSGGDPIFEGWMRSTESSYKDKFRGMVGFSPVSHRITAGCDILLMPSSRFEP 540
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
Db 483 redvqfvmlysggdpelcdwmrstesifkdkfrgwgfsvpvshritagcdillmpsfef 542
QY 541 CGLNQLYAMQYGVVPHVHGTGGLRDTVETENPFCAGKEEGTGNFAFSLPTVDKMLWALRTA 600
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
Db 543 cglinqlyamqygtvpvvhggtggrldtvetfnpgakgeegtgwafagpltdvdkmlwalrta 602
QY 601 MSTFREHKPSWEGLMKRGMTKDHWDHA 628
||| : ||| || : || : || : || : || : || : || : || : || : || : || : ||
Db 603 istyrehkssweglmkrmtkdhtwdha 630

RESULT 5
AAW51231
ID AAW51231 standard; Protein: 626 AA.
XX
XX AC AAW51231;
XX
XX DT 16-NOV-1994 (first entry)
XX
XX DE Soluble rice starch synthetic enzyme.
XX
XX KW Rice; starch synthetic enzyme; transit peptide; amyloplast;
XX transition; protoplast; expression.
XX
XX OS Oryza sativa.
XX
XX FH Key Location/Qualifiers
FT Active-site 1..113 /label= transit_peptide
FT /note= "claim 3"
FT Protein 114..626 /label= mat_protein
FT /note= "claim 1"
FT
XX
XX PN JP06070779-A.
XX
XX PD 15-MAR-1994.
XX
XX PF 07-JUL-1992; 92JP-0179947.
XX
XX PR 07-JUL-1992; 92JP-0179947.
XX
XX PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
XX
XX DR WPI: 1994-128678/16.
XX DR N-PSDB; AAQ45183.
XX
XX PT Soluble rice starch synthetic enzyme gene and transit peptide -
PT for the efficient transport of heterologous proteins to
```











```

XX CC The sequence is that of maize starch synthase from pEXS52.
CC CC It can be used in the production of a hybrid polypeptide
CC CC comprising a starch-encapsulating region (SER) fused
CC CC to a payload protein. The hybrid polypeptide can be used to make
CC CC modified starches comprising the payload protein, selected from,
CC CC e.g. hormones, growth factors, antibodies, enzymes, dyes,
CC CC immunoglobulins, etc. The modified starch can also be used
CC CC to provide grain feeds enriched in amino acids. By encapsulating
CC CC the payload protein in starch, it is more resistant to
CC CC degradation by stomach acids.
XX CC
SQ Sequence 539 AA;

Query Match 59.1%; Score 2390.5; DB 19; Length 539;
Best Local Similarity 88.4%; Pred. No. 2.4e-196;
Matches 444; Conservative 23; Mismatches 32; Indels 3; Gaps 2;

QY 95 VGELA-PDLLL--EGIAEDSIDSIIVAASEQDSEIMDANEQQAQVTRSIIVFTGEAAPY 151
Db 2 vaelrsredlglepegiaeagldntvvvaseqdeivvgkeqarakvtqsvftgeaspy 61
QY 152 AKSGGLGDCGSLPIALAAARHRVWVMPRYLNGSSDKNTAKALYTAHHKIKPCFGSGHE 211
Db 62 aks99lgdcvgsipvalaarghrvmvmprylngtsdknyanafytekhiripcfggehe 121
QY 212 VTFEYHRDNDVWVFDHPHSYHRPGSLYGDNFGAFGDNQFRYTLCCVAAACEAPLILELGG 271
Db 122 vtfefeyrdsdvwfvfdhpsyhrpgnlygdkfgafgdnqfyllcyaaaceaplilel 181
QY 272 YIYGQNCMFVNDWHASLVPLVLAARYPYGVYRDRSRTLVIHNLHQGVPEASTYPDLG 331
Db 182 yiy9qncmfvndwhaslvplvllaarypygvkykdsrslivihnlhahqgvpeastypdl 241
QY 332 LPPEWYGALEWPEWARRHALDKGEAVNFKGAVVTADRIVTVSOGYSWEVTTAGGQG 391
Db 242 lppewgaalewfpewarrhaldkgeavnfkagavvtadrivtvsgyswevttaggqg 301
QY 392 LNELLSRKSVLNGVINGIDINDWNPSTDKCLPHYSVDDLSGKAKCAELQKELGLPVR 451
Db 302 lnellsrksvlnvgidindwnpstdkciphysvddlsqkckgalkelglpir 361
QY 452 EDVPLIGFGRLDYQKQIDLIKMAIPELMRDVOFVHLGSGDPIFCGWMRSTESSYKDKF 511
Db 362 pdvpllgfgrldyqkgidliqilpdlmrdevqfvmhgsgdpelcdwmrstesifkdkf 421
QY 512 RGWGFSPVSHRITAGCDITLLMPSRFEPCGGLNOLYAMOYGTVPVHGTGGLRDTVETN 571
Db 422 rgwgfspvshritagcdillmrsfepcglnqlyamdygtvpvvhacggrldrtvenfn 481
QY 572 PFGAKGEGTGWAFSPLTVDKM 593
Db 482 pfgekgegtgwafaplttenm 503

RESULT 9
AAR99540
ID AAR99540 standard; Protein; 459 AA.
XX
AC AAR99540;
XX
XX 31-OCT-1996 (first entry)
XX
XX Soluble starch synthase (Isoform B).
XX
XX Soluble starch synthase; transformation; transgenic plant; starch;
KW amylose; amylopectin.
XX
XX Solanum tuberosum.
XX
XX DE4441408-AL.
XX
```

```

PD 15-MAY-1996.
XX
PF 10-NOV-1994; 94DE-4441408.
XX
PR 10-NOV-1994; 94DE-4441408.
XX
PA (GENB-) INST GENBIOLOGISCHE FORSCHUNG.
XX
PI Abel GJ, Kossmann J, Springer F;
XX
DR WPI; 1996-240218/25.
XX
DR N-PSDB; AAT32325.
XX
PT DNA encoding soluble starch synthase of potato - used to produce
PT transgenic plants with increased produ. of starch or able to produce
PT modified starches
XX
PS Claim 11; Page 20-22; 32pp; German.
XX
XX DNA sequences from the potato encoding soluble starch synthase
CC having this 459 C-terminal amino acid sequence or the 677
CC C-terminal amino acid sequence given in AAR99539 can be used to
CC identify and isolate homologous sequences encoding soluble starch
CC synthase and enzymes with similar activities from plants or other
CC organisms; to transform prokaryotic or eukaryotic cells; to produce
CC transgenic plants which synthesise starch of altered structure or
CC in increased yield.
XX
SQ Sequence 459 AA;

Query Match 42.5%; Score 1718.5; DB 17; Length 459;
Best Local Similarity 70.6%; Pred. No. 8.4e-139;
Matches 303; Conservative 62; Mismatches 61; Indels 3; Gaps 2;

QY 203 IPCFGSHEVTFEYHRDNDVWVFDHPHSYHRPGSLYGDNFGAFGDNQFRYTLCCVAAACE 262
Db 14 vhcfdgaqevafyheyragdvwfvdhssyrrpgdygafgdnqfyllshaace 73
QY 263 APLILELGGYIYGONCMFVNDWHASLVPLVLAARYPYGVYRDRSRTLVIHNLHQGV 322
Db 74 aplvlpigftygekelcflandcnaalvplllaarypygvkykdsrslivihnlhahqgve 133
QY 323 PASTYPDGLPPEWYGALEWPEWARRHALDKGEAVNFKGAVVTADRIVTVSOGYSWE 382
Db 134 pavtynnlgppqwvgavewifptwarahaldtgecnvfkgaiaavadrlitvsggyswe 193
QY 383 VTTAGGGGGLNELLSRKSVLNGVINGIDINDWNPSTDKCLPHYSVDDLS--GKAKCKA 440
Db 194 ittpggyglhellssrqslnvgitngidvndwnpstdehshysindlspgkvqckt 253
QY 441 ELQKELGLPVREDVPLIGFGRLDYQKQIDLIKMAIPELMRDVOFVHLGSGDPIFCGWM 500
Db 254 dlqkelglpvrpdcpligfgrldyqkgvdiilsaipelmqndvqvmvmlgsgekqdw 313
QY 501 RSTESSYKDKRFGWGFSPVSHRITAGCDITLLMPSRFEPCGGLNOLYAMOYGTVPVHGT 560
Db 314 rhtenlfkdkfrawgvfnvpsvshritagcdillmrsfepcglnqlyamdygtvpv 373
QY 561 GGLRDTVETNPFQAKGE--EGTGWAFSPLTVDKMALMALRTAMSTPREHKPSWEGLMKRCM 619
Db 374 ggirdtvkdfnpaqegkgegtwtgwtfspltseklftiklairtytehksksweglmkrgm 433
QY 620 TKDHTWDHA 628
Db 434 grdyswena 442

RESULT 10
AAW56487
ID AAW56487 standard; Protein; 698 AA.
XX
AC AAW56487;
```



```
XX 11-SEP-1998 (first entry)
DT Zea mays soluble starch synthase IIB.
XX
XX SER; starch-encapsulating region; fusion vector;
KW soluble starch synthase IIB; glucosyl transferase.
KW
XX Zea mays.
XX
XX W09814601-A1.
PN
XX 09-APR-1998.
XX
XX 30-SEP-1997; 97WO-US17555.
XX
XX 30-SEP-1996; 96US-0026855.
XX
XX (EXSE-) EXSEED GENETICS LLC.
XX
XX Guan H, Keeling P;
XX
XX WPI; 1998-240100/21.
DR
XX N-PSDB; AAV29755.
DR
XX
XX Hybrid polypeptide comprising starch-encapsulating region and
PT protein - useful for, e.g. producing protein(s) resistant to
PT degradation by stomach acids
XX
XX Example 2; Page 36-38; 156pp; English.
PS
XX The sequence is that of soluble starch synthase IIB.
XX
XX It can be used in the production of a hybrid polypeptide
CC comprising a starch-encapsulating region (SER) fused
CC to a payload protein. The hybrid polypeptide can be used to make
CC modified starches comprising the payload protein, selected from,
CC e.g. hormones, growth factors, antibodies, enzymes, dyes,
CC immunoglobulins, etc. The modified starch can also be used
CC to provide grain feeds enriched in amino acids. By encapsulating
CC the payload protein in starch, it is more resistant to
CC degradation by stomach acids.
XX
XX Sequence 698 AA;
SQ
Query Match 29.8%; Score 1205.5; DB 19; Length 698;
Best Local Similarity 43.0%; Pred. No. 1.7e-94;
Matches 283; Conservative 83; Mismatches 221; Indels 71; Gaps 17;
QY 12 APSVRLRADPA-----TAARASACVVRRLRLRLRGRVVAELSRGPAARPAQQQLAPP 66
DB 56 aasvraaaapaggeseeaaakssss-sqagavvggstakavdsasppnpltsapkgsqsaam 114
QY 67 L--VPGFLAPPPPAQSPAQTQPLP-----DAG-----VGLAPDLLLEGTAEDSI 112
DB 115 qngtsgsssaastaapvqpkadhpaavtkreidasavkpepagddarpvesigiaepv 173
QY 113 DSIIIVAASEQDS-----EIMDANEQPOA-KVTRISIVFTGEAAPYAKSGGLGDVCGSLP 165
DB 174 dakadaapatdaaasapydrdnepgplagnvnmvvvvvasecapfctgtgldvvgalp 233
QY 166 TALAARHVRVVMVPRYLNGSSDKNYAKALYAKIKIPCGGSHEVTFRHEYRDNVDWV 225
DB 234 kalargrhvmvlpvry-----geyaearldlgvrrrykvagqdvsevfysyldgvdv 287
QY 226 FVDHPSY--HRPGSLYGDNFGAFDGNQFRYTLICYAACBAPLILELGGYIYGQ--NCMFYVN 283
DB 288 fveapfrrhnniyy---gerldilkmlfckaavevpyapcggtvygdgnlvfian 344
QY 284 DHASLVPVLLAARYPYGVYDRSSTLVTHNLHQGVPEASTYPDLGLPPEWYGALEWV 343
DB 345 dwhtallpylkayyrnglmgqarsvlihviahqgrgvpvddfvnfldp----- 394
```

```
QY 344 FPEWARRHALDK-----GEAVNELKCAVVTADRIVTVSOGYSHEVTTAEGGGLNEL 395
DB 395 -----ehyldhfklydniggdhsnvfaaglkktadrvvtvsngymwelktseggwglhdi 448
QY 396 LSSRSKSVLNGIYNGIDINDWNPPTDKCLPH---HYSVDL--SGRAKCKAEQLKELGLPV 450
DB 449 inqndwkllgglvngldmswnpavdvhlhsddytnyfcldtqkrqckaalgqlglqv 508
QY 451 REDVPLIGFTGRIDYQKIGDLIKMAIPELMREDVQFVMIIGSDPIFEGWMRSTESSYKDK 510
DB 509 rddvpligfgrldhckgvdliladaihwiagqdvqlvmlgtgradledmlrrfesehsdk 568
QY 511 FRGWGFSVPVSHRITAGCDILLMPSPFPCGLNQLYAMQYGTVPVVGCTGGRLDHFVTF 570
DB 569 vrawvgfsvplahrifitagadillmpsfepcglinqlyamaygtvpvvhavggldcvapf 628
QY 571 NPEGAKGEGCTGWAFFPLTVDKMLWALRTAMSTFRHKKPSWEGLMKRGMTKDHTWDHA 628
DB 629 dpf---ndtglwtfdraeanrmidalschlttyrnykeswacracrgmaedlswdha 683
RESULT 11
AAW70892
ID AAW70892 standard; Protein; 804 AA.
XX
AC AAW70892;
XX
DT 23-AUG-1999 (first entry)
XX
DE Maize starch soluble synthase IIB.
XX
KW Non-glycogen-like polysaccharide production; fermentation;
KW starch synthesis enzyme; glycogen-synthesis enzyme; glycogen synthesis;
KW non-starch branching gene; amylopectin; amylose; plant-like starch;
KW maize starch soluble synthase IIB.
XX
OS Zea mays.
XX
PN W09844780-A1.
XX
PD 15-OCT-1998.
XX
PF 03-APR-1998; 98WO-US06660.
XX
PR 04-APR-1997; 97US-0042939.
XX
PA (EXSE-) EXSEED GENETICS LLC.
XX
PI Guan H, Keeling PL;
XX
DR WPI; 1998-568285/48.
XX
DR N-PSDB; AAV70958.
XX
PT Producing non-glycogen-like polysaccharides in bacteria, fungi or
PT plants - transformed with genes for enzymes involved in starch or
PT glycogen synthesis allows fermentative production of starches with
PT engineered properties
XX
XX Disclosure; Fig 50; 150pp; English.
XX
XX The specification describes a method for the production of
XX non-glycogen-like polysaccharides in a host. The method comprises
XX transforming a host, suitable for fermentation, with genes encoding
XX starch- or glycogen-synthesis enzymes, and fermenting the transformants.
XX The specification also describes hosts transformed with a gene active
XX in glycogen synthesis and at least one non-starch branching gene,
XX involved in production of amylopectin or amylose in its original host.
XX The method is used to produce plant-like starches by fermentation and
XX new starches in plants. These starches are useful for all food and
XX non-food applications of starch. The present sequence is used in
XX the course of the invention.
XX
XX Sequence 804 AA;
SQ
```



```

Query Match      29.8%; Score 1203.5; DB 19; Length 804;
Best Local Similarity 41.7%; Pred. No. 3.2e-94;
Matches 287; Conservative 85; Mismatches 229; Indels 87; Gaps 18;

QY 12 APSVRLRADPA-----TAARASACVVRRLRLARGRYVAELSRGPAARPAQOQLAPP 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56 aasvraaapagseaeaksss-sqagavqgstakavdsasppnpltsapksasaam 114
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 67 L--VPGFLAPPAPPAQSPATQPLP-----DAG-----VCELAPDLLLEGIAEDSI 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 qngtsgsesaastapagpkadhpsapvtkreidasavkpepadgdarpviesiglae-pv 173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 113 DSIIVAASEQDS-----EINDANEQQA-KVTRSIIVFTGEAAAPYAKSGGLGDDVCCSLP 165
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174 dakadaapatdaasapydrednepglagpnvmvvaasecapfctgglgavgalp 233
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 166 IALAARGHRVVMVPRYLNGSSDKNYAKALYTAHKIKIPCGGSHVETFFHEYRDNDVW 225
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 kalarrghrmvmpvpry-----geyaearldlgvrrrykvagqdsenvtyhsyldgvdfv 287
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 226 FVDHPST-HRPGSLYGNFGAFGNQPRYTLILCYAAACEAPLILELGGYIYGO-NCMFVNV 293
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 fveappfrhcnhiyg--geridilkrmilfckaavevpwypagcgvvgdgnlvfian 344
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 284 DWHASLVPVLLAAKARYPYGVYRDSRSTLVTHNLAHQGVPEASTYPDLGLPPEWYGALEWV 343
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 345 dwhallpvyllkayrdnglmqyarsvlvlniahgqrgpvddfvnfdlp----- 394
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 344 FPEWARRHALDK-----GEAVNFKGAVVTADRTVTVSQGSWEVTTAEGQGLNEL 395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 395 -----ehyldhfklydnigdhshnvfaagtktdrvvtvsnymwelktseggwghdi 448
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 396 LSSRSKSLVINGIDINDWNPNTDKCLPH-----HYSVDDL-SGAKCKAELOKELGPV 450
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 inqndwklqgvingidmewnpavdvhllshddytntfctldgkrqckaqlrglqlv 508
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 451 REDVPLIGFTRLDYKGDILKMAPELMDREDVQFVMLGSGDPFEGWMRSTESSYKOK 510
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 rddvpligfigrldhqdqdiadahlwtagdqvqlvmlgtgradledmlrrfesehsdk 568
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 511 FRGWGFSVPVSHRITAGCDILLMPSEPCGLNQLYAMQYGVTVVHGTGGLRDTVETFF 570
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 569 vrawgfsvplahritagadillmpsfepcgnlqlyamaygtvvpvhavggldrtvapf 628
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 571 NPGKAGEECTGWAFSPLTVDKMLALRTAMSTFREHKPSWEGLMKRGMTKDHTWDHAP- 629
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 629 dpf---ndtqlgwtfdraeanrmidalschltttrnykeswacragmaedlswdhaav 685
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 630 -----SSTRSSSGPSWT 642
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 686 lyedvlvkakyqvanlatrrscrtwt 713
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
AAB37567
ID AAB37567 standard; Protein; 799 AA.
XX
AC AAB37567;
XX
XX
DT 01-MAR-2001 (first entry)
XX
DE Wheat starch synthase II SEQ ID NO: 4.
XX
KW Wheat; starch synthase; SSII; starch content; starch synthesis;
KW food product; adhesive.
XX
OS Triticum aestivum.
XX
PN WO200066745-A1.
XX
PD 09-NOV-2000.

```

```

XX PF 28-APR-2000; 2000WO-AU00385.
XX PR 29-APR-1999; 99AU-0000052.
XX XX
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA (GOOD-) GOODMAN FIELDER LTD.
PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
XX
XX Morell M, Li Z, Rahman S, Appels R;
XX WPI: 2000-647602/62.
XX DR N-PSDB; AAC66411.
XX
XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and
PT WST-II, useful in modifying plant starch content and/or composition -
XX
XX Claim 19; Page 158-161; 21pp; English.
XX
XX The present invention relates to novel protein and coding sequences from
CC wheat. The proteins are wheat starch synthases, designated SSII and
CC SSIII. These can be used in the modification of plant starch content or
CC composition, and to screen plants to identify mutations which affect
CC starch content and composition. The starch can then be used in food
CC products, such as flour, and in films, coatings, adhesives, building
CC materials and packaging materials.
XX
XX Sequence 799 AA;

Query Match      29.0%; Score 1174.5; DB 21; Length 799;
Best Local Similarity 37.1%; Pred. No. 9.8e-92;
Matches 286; Conservative 91; Mismatches 223; Indels 171; Gaps 19;

QY 7 GAGCLA-----PSVRLRADPATAARA-----SACVVRRLRLARGRYVAE----- 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36 gagrhlwppwppqtrtdagdvaaagkdkdarvdddaasargprarrggaatkvaerdp 95
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 48 ---LSPE-----GPAARPAQOQLAPPPLVPGFLAPP-----PAPAQS 82
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 vktldraaegapappaprpqdaarp--psmngtpvngenkstgggatkdsghlpapara 153
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 83 PAP-TQPLPADAGVGEL-----APDLLLEGIAEDSIDSIIVAASEODSET----- 126
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 154 phpstqnrpvngenkvanvasptsiaevvapsaatisiskapesvvpaeekpppsgs 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 127 -----MDANEQPAKV----- 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 nfvsasaprldidsdvepelkkgaviveeapnkalppaavqgdldwfkkyifgee 273
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 138 -----TRSIIVFTGEAAAPYAKSGGLGDDVCCSL 164
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 pveakddgwavaddagsfehhqnhdsgplagenmnmvvaacsckwtgglgavagal 333
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 165 PIALAARGHRVVMVPRYLNGSSDKNYAKALYTAHKIKIPCGGSHVETFFHEYRDNDVW 224
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 pkalakrghrmvmpvpry-----gdyeeaydvgrkryk--aagqdmvnyfnayldgvdf 387
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 225 VFVDHPST-HRPGSLYGNFGAFGNQPRYTLILCYAAACEAPLILELGGYIYGO-NCMFVNV 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 388 vfidaplfrhrqediyyg---gsrqeimkrmilfckaavevpwypcgvvgdgnlvfia 444
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 283 NDWHASLVPVLLAAKARYPYGVYRDSRSTLVTHNLAHQGVPEASTYPDLGLPPEWYGALEWV 342
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 ndwhtallpvyllkayrdnglmqytrslmvlhiahgqrgpvdefpfelpehyl----- 499
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 343 VPEWARRALDKGEAVNFKGAVVTADRTVTVSQGSWEVTTAEGQGLNELSSRSKS 402
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 500 ---ehfrlydpvggehanyfaagkmadqvvvvpspgylwelktveggwghdiirqndwk 556
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 403 LNVGVNGIDINDWNPNTDKCLPH-----HYSVDDL-SGAKCKAELOKELGPVREDVPLI 457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 557 trgivngidmewnpdvdhksdgytnfsigtldsgkrqckealgrgelglqvradvp11 616
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



QY 458 GFTGRDYQKIDILKMAIPELMREDVQFVMLGSDPIFEGWNRSTESSYKDKFRGWGCF 517  
 DB 617 gfigridgkgveiadampwivsgdqvilmigtgrhdielsmlrhferehdkvrgwvdf 676  
 QY 518 SVPVSHRITAGCDILLMPSEPCGLNOLYAMOYGTVPVVGHTGGURDTVETFNPFKAGK 577  
 DB 677 svrlahritagadallmpsfepcgnlqlyamaygtvpvhavggvrdtvpfpdpl---n 733  
 QY 578 EEGTGNAFSPLTVDKMLWALRTAMSTFRHKPSWEGMLKRGMTKDHWTWHA 628  
 DB 734 hsglgtwfdraeahkkliealghclrtvdykeswrglqergmsqdfswaha 784

RESULT 13  
 AAB37597  
 ID AAB37597 standard; Protein; 799 AA.  
 AC AAB37597;  
 XX 01-MAR-2001 (first entry)  
 DT Wheat starch synthase II protein.  
 DE Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;  
 KW food product; adhesive.  
 XX Triticum aestivum.  
 OS WO200066745-A1.  
 PN 09-NOV-2000.  
 PD 28-APR-2000; 2000WO-AU00385.  
 PF 29-APR-1999; 99AU-0000052.  
 PR (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
 PA (GOOD-) GOODMAN FIELDER LTD.  
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.  
 XX Morell M, Li Z, Rahman S, Appels R;  
 PI WPI; 2000-647602/62.  
 DR Nucleic acid molecules encoding wheat starch synthase (WST)-I and  
 XX WST-II, useful in modifying plant starch content and/or composition -  
 PT Example 9; Fig 3; 21pp; English.

The present invention relates to novel protein and coding sequences from wheat. The proteins are wheat starch synthases, designated SSII and SSIII. These can be used in the modification of plant starch content or composition, and to screen plants to identify mutations which affect starch content and composition. The starch can then be used in food products, such as flour, and in films, coatings, adhesives, building materials and packaging materials.

Sequence 799 AA;

Query Match 29.0%; Score 1174.5; DB 21; Length 799;  
 Best Local Similarity 37.1%; Pred. No. 9.8e-92;  
 Matches 286; Conservative 91; Mismatches 223; Indels 171; Gaps 19;

QY 7 GAGCLA---PSVRLRAPATAARA-----SACVVRARLRLRARGRYVAE---- 47  
 DB 36 gagrlhwppqrtdggvaaragkdkardvdddaasarprrarrgaatkvaerdp 95  
 QY 48 ---LSRE----GPAARPAQOQOLAPPLVPGFLAPPP-----PAPAQS 82  
 DB 96 vktldrdaaeggapapaprqdaarp--psmngtptvngenkstgggatkdgslpapara 153

QY 83 PAP-TQPPLPDAGVGEL-----APDLLLEGIAEDSDISITVAASEQDSSEI----- 126  
 DB 154 phpsctqrvpvngekanvaspptsiaevvapdsaatissdkapesvvpaepppsags 213  
 QY 127 -----MDANEQPOAKV----- 137  
 DB 214 nfvsaasaprlidisdvpeklkkgaviveeapnkalappaavqedlwkfyigfee 273  
 QY 138 -----TRSIIVFTGEAAAPYAKSGGLGDCVCSL 164  
 DB 274 pveakddgavavaddagsfehhqnhdsgplagenvmnmvvaecspwcktaglgdvagal 333  
 QY 165 PIALAARGHRVMVMPRYLNGSSDKNYAKALYAKHIIKIPCGGSHETTFHEHYEDNDWM 224  
 DB 334 pkalakrghrvmvvpvry---gdyeeaydvgrkyk--aagqdmvnyfhayldgvdf 387  
 QY 225 VFVDHPSY-HRPGSLYGDNFGAFGDNQFRYTLILCYAAACEAPLILELGVYIGQ-NCMFVV 282  
 DB 388 vfideplfrhrqediyg---gsrqelmkfmllfckaaevpwhvpcggyvpydgdnlvfla 444  
 QY 283 NDWHASLVPVLLAAKYRPGYVRDSRSTLVIIHNLHQGVPEASTYPDGLGPPEWYGALEM 342  
 DB 445 ndwhtallpylkayrhdglmqytrsimvlniahqgrgpyvdefptelpehyl----- 499  
 QY 343 VPEWARRHALDKGEAVNFKGAVVTADRIIVTSQGYSEVTTAEGGQGLNELLSRSKSV 402  
 DB 500 ---ehfrlydpvggehanyfaaglkmadgvvsvpylweilktvaggwglhdiirndwk 556  
 QY 403 LNCIVNGIDINDWNTPTDKCLPH-----HYSVDDL-SGKAKCAEKLQKELGLPVREDVPLI 457  
 DB 557 trglvngldnmewnpevdvhlksdgytnfsldsgkrcqkealqrelglqvradvpil 616  
 QY 458 GFTGRDYQKIDILKMAIPELMREDVQFVMLGSDPIFEGWNRSTESSYKDKFRGWGCF 517  
 DB 617 gfigridgkgveiadampwivsgdqvilmigtgrhdielsmlrhferehdkvrgwvdf 676  
 QY 518 SVPVSHRITAGCDILLMPSEPCGLNOLYAMOYGTVPVVGHTGGURDTVETFNPFKAGK 577  
 DB 677 svrlahritagadallmpsfepcgnlqlyamaygtvpvhavggvrdtvpfpdpl---n 733  
 QY 578 EEGTGNAFSPLTVDKMLWALRTAMSTFRHKPSWEGMLKRGMTKDHWTWHA 628  
 DB 734 hsglgtwfdraeahkkliealghclrtvdykeswrglqergmsqdfswaha 784

RESULT 14  
 AAB37566  
 ID AAB37566 standard; Protein; 798 AA.  
 AC AAB37566;  
 XX 01-MAR-2001 (first entry)  
 DT Wheat starch synthase II SEQ ID NO: 2.  
 DE Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;  
 KW food product; adhesive.  
 XX Triticum aestivum.  
 OS WO200066745-A1.  
 PN 09-NOV-2000.  
 PD 28-APR-2000; 2000WO-AU00385.  
 PF 29-APR-1999; 99AU-0000052.  
 PR (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
 PA (GOOD-) GOODMAN FIELDER LTD.  
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.  
 XX Morell M, Li Z, Rahman S, Appels R;  
 PI



XX WPI: 2000-647602/62.  
DR N-PSDB; AAC86410.  
XX  
PT Nucleic acid molecules encoding wheat starch synthase (WST)-I and  
PT WST-II, useful in modifying plant starch content and/or composition -  
XX  
PS Claim 19; Page 152-155; 21pp; English.  
XX  
CC The present invention relates to novel protein and coding sequences from  
CC wheat. The proteins are wheat starch synthases, designated SSII and  
CC SSIII. These can be used in the modification of plant starch content or  
CC composition, and to screen plants to identify mutations which affect  
CC starch content and composition. The starch can then be used in food  
CC products, such as flour, and in films, coatings, adhesives, building  
CC materials and packaging materials.  
XX  
SQ Sequence 798 AA;  
  
Query Match 28.8%; Score 1165; DB 21; Length 798;  
Best Local Similarity 37.4%; Pred. No. 6.4e-91;  
Matches 288; Conservative 90; Mismatches 222; Indels 170; Gaps 22;  
  
QY 7 GAGCLA-----APSVRLRADPATARASA-----CVRRLRLRLARG-----RYVAE----- 47  
DB 36 gagrlhwppbpqrtardgavaaraagkddaglddaapargalrggaatkvaerrdpv 95  
QY 48 --LSRE----GPAARPAQOQOLAPPLVPGFLAPP-----PAPAQSP 83  
DB 96 klldraaeggaspapqrqeda--rlpsmngmpvngenkstgggatkdsqglpaparp 153  
QY 84 APT-----QPP-----LPD-----AGVCELAPDLL--EGIAEDSID 114  
DB 154 qpssqrvpvgengkanvaspptsiaevaapdaatissidkapesvpaekappssgn 213  
QY 115 IIVASEQDSEIMD-----ANEQPAKV----- 137  
DB 214 fypsasagdsdvsdvelelkgavivkeapnkalpspaapavqgdlwdfkkyigfeep 273  
QY 138 -----TRSTVFTVGEAAPYAKSGGLGDVCGSLP 165  
DB 274 veakddgravadagsfehhqhdsgplagenmvmvvaacspwcktgglgdvagalp 333  
QY 166 IALAARGHRVMVPRYLNSSDKNYAKALYTAHKIKPCFGSGHEVTFHEYRONVDWV 225  
DB 334 kalakgrhvmvpry----gdyeeaydvgrkyk---aaqdmeyvfhayidgvdiv 387  
QY 226 FVDHFSY--HRPGSLYGDNFGAGDNQFRYLLCYAACEAPLILELGGYTYGO--NCMFVYN 283  
DB 388 fidaplfrhrqediyy---gsrqelmkrmlfckaavevphwpcgvpvgdgnlvfian 444  
QY 284 DWHASLVPLLAAKVRPGYGVYRDSRSTLVIHNLAHQGVPEASTYDGLGLPPWYGALEWV 343  
DB 445 dwhtalpylkaayrdhgmtyrsimvhlahqgrpvdeftpelpehyl----- 498  
QY 344 FPEARRHALDGEAVNFKGAVTADRTVTSQGYSWEVTTAEGGQGLNELLSRKSVL 403  
DB 499 --ehfrlydpvggehanyfaaglkmadqvsvpspylwlkvtvegwgldiirqndwt 556  
QY 404 NGIVNGIDINDNPTDKCLPH-----HYSVDL--SGKAKCKAELQELGLPVREDVPLTG 458  
DB 557 rgivngidmewnpvvdhalksdgytnfslgtldsgkrqckeaqlrelglvrgdvpllg 616  
QY 459 FTRGLDYOKGIDLIKMAIPELREDVQFVMLGSGDPIFEGWMRSTESSYKDKFRGWGVS 518  
DB 617 figrlgdkqgveiidampwivsqdvqlvmlgtrghdlegmlrhferdhdkvrgwvgs 676  
QY 519 VPVSHRITAGCDIILLMPSPFEPCGLNQLYAMQYGVVPHVGTGGLRDTVETFPNFGAKE 578  
DB 677 vrlahritagadallmpsrfeqcnlqyamaygtvppvhavgrldrtvppfdpf---nh 733  
QY 579 EGTGNWAFSLTVDKMLWALRTAMSTFREHKPSWEGMLKRGMTKDHTWDHA 628

DB 734 sglgwtfraeaqklliealghclrtyrdykeswrglqergmsqdfswuha 783  
RESULT 15  
AAW23938  
ID AAW23938 standard; Protein; 799 AA.  
XX  
AC AAW23938;  
XX  
DT 21-MAY-1998 (first entry)  
XX  
DE Wheat granule-bound starch synthase.  
XX  
KW Starch synthase; wheat; transgenic plant.  
XX  
OS Triticum aestivum L. cv. Florida.  
XX  
PN WO9745545-Al.  
XX  
PD 04-DEC-1997.  
XX  
PF 28-MAY-1997; 97WO-EP02793.  
XX  
PR 11-SEP-1996; 96DE-1036917.  
PR 29-MAY-1996; 96DE-1021588.  
XX  
PA (AGRE ) HOECHST-SCHERING AGREVO GMBH.  
XX  
PI Block M, Loerz H, Luetticke S, Froberg C, Kossmann J;  
PI Walter L;  
XX  
DR WPI: 1998-032652/03.  
DR N-PSDB; AAV01528.  
XX  
PT Nucleic acid encoding starch synthase enzymes from wheat - for  
PT transgenic plants that produce modified forms of starch, useful e.g.  
PT in foods, or for production of packaging materials and disposable  
PT goods  
XX  
PS Claim 1; Page 54-58; 71pp; English.  
XX  
CC This amino acid sequence comprises a full-length sequence for a  
CC granule-bound starch synthase of summer wheat (cv. Florida). It  
CC was deduced from a cDNA clone (AAV01528) isolated from a 21-day  
CC carypopsis cDNA library. A solubled starch synthase (see AAW23937)  
CC has also been identified. Isolated nucleic acids encoding these  
CC enzymes can be inserted into vectors for production of transgenic  
CC plants, particularly starch-producing plants, specifically wheat.  
CC Use of the isolated nucleic acids, or of antisense sequences, allows  
CC starch metabolism to be regulated in transgenic plants.  
CC Overexpression may result in improved crop yield, while modification  
CC of starch in planta may eliminate the need for subsequent  
CC chemical/physical modification. Plants with altered levels of the  
CC various isoforms of starch synthase will produce starch of different  
CC chain length, amylose/amylopectin ratio, degree of branching,  
CC phosphate content, gelatinisation behaviour, granule size and shape,  
CC viscosity etc. The starch produced by such plants is useful  
CC particularly in foods (especially bakery goods and pasta) or to  
CC produce packaging materials or disposable goods, as well as in any  
CC other known use of starch.  
XX  
SQ Sequence 799 AA;  
  
Query Match 28.4%; Score 1148.5; DB 19; Length 799;  
Best Local Similarity 37.0%; Pred. No. 1.7e-89;  
Matches 286; Conservative 86; Mismatches 226; Indels 175; Gaps 20;  
  
QY 7 GAGCLA-----PSVRLRADPATAA-----RASACVVRRLRLARGRYVAE----- 47  
DB 36 gagrlhwppbpqrtardgavaalaagkddaglddaasvrgqralrggaatkvaerrdp 95



```

QY 48 ---LSRE-----GPAARPAQOQOLAPLPVPGFLAPP-----PAPA 80
Db 96 vktldraaeggpsppaarppsmngm-----pvngenkstgggatkdsigtpa 151
QY 81 QSPAP-TQPPLPDAGVCEL---APDLLLEGIAEDSIDSIIV----- 117
Db 152 rapbstqnravngenkanvaspptsiaaaaaasdaatisidkapesvvpaektpss 211
QY 118 -----AASQDSEIMDANEQ-----PQAKV----- 137
Db 212 gsnfesaasagdsdvsdeqelkgavvveeapkpalsappaapavqedlwfkkkyigf 271
QY 138 -----TRSIIVFTGEAAPYAKSGGLGDCVCG 162
Db 272 eepveakddgravadagsfehhqndsgplagenvmnvvvaecspwcktgglgdvag 331
QY 163 SLPIALAAARGHRVVMVPRYLNSSDNKNAKALYTAKHIPICFGGSHEVTFEHEYRDNV 222
Db 332 alpkaakrghrvmvvpry---gdyeeaydvgrkyk--aagqdmvnyfhayidgv 385
QY 223 DWYFVDHPSY-HRPGSLYGNFAGFQDNOPRYTLLCYACEAPLIILELGGYIYQ-NCMF 280
Db 386 dfvfidaplfrhrqediyg--gsrqeimkrmilfckaaavevpwhvpcggvpygdgnlvf 442
QY 281 VVNDWHASLPVLLAAKYRPYGVYRDSRSTLVITHLAHQGVBPASTYPDLGLPPEWYGAL 340
Db 443 iandwtallpvylykayrddhgmqytrsimvlniahqgrgvpvdefpfelpehyl--- 499
QY 341 EWVFPEWARRALDKGEAVNFKGAVVTADRIVTVSQYSWEVTTAEGGGLNELSSRK 400
Db 500 -----ehfrlydpvggehanyfaaglkmadqvvvvsgpylweiktvegglhdiirqnd 554
QY 401 SVLNGIUNGIDINDWNPTTDKCLPH---HYSVDDL-SGKAKCAELQKELGLPVREDVP 455
Db 555 wktgivingidnmewnpvcdahiksdgynfsirtidsgrkqckealqrelglqvradvp 614
QY 456 LIGFGRLDYQKIDLIKMAIPELMRDQFVMLGSDPIFEGWMRSTESSYKDKPRGWV 515
Db 615 llgfrldgqkgveiliadampwivsgdqlvmlgtgrhlesmlqhferrehdkvrgwv 674
QY 516 GFSVPVSHRITAGCDILLMPSRPEPCGLNQLYAMQYGTVPVHGTGGLRDTVETFPFGA 575
Db 675 gfsvriahritagadalmpsrfeqpcglnglyamaygtvpvhavggldrtvppfdpf-- 732
QY 576 KGEEGTGWAFSPLTVDKMLMALMTAMSTFREHKPSWEGMLMKRGMTKDHWTWDAH 628
Db 733 -nhsglgwtfdraeahklliealghcirtyrdfkeswralqergmsqdfswaha 784

```

Search completed: March 28, 2002, 16:19:28  
Job time: 5033 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2002, 16:19:46 ; Search time 78.51 Seconds  
(without alignments)  
669.358 Million cell updates/sec

Title: US-09-674-824-2  
Perfect score: 4044  
Sequence: 1 MAATGVGAGCLAPSVRLRAD.....SDGSLSVRTAEIRNQLVTL 756

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 335014 seqs, 69512255 residues

Total number of hits satisfying chosen parameters: 335014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA\_New.\*

1: /cgn2.6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2.6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2.6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2.6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2.6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2.6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
7: /cgn2.6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
8: /cgn2.6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4044	100.0	756	5	US-09-674-824-2
2	3612	89.3	671	6	US-09-952-677-2
3	3348	82.8	647	6	US-09-508-377-14
4	2749	68.0	649	6	US-09-931-297-2
5	1196.5	29.6	801	7	US-10-044-543-26
6	1174.5	29.0	799	7	US-10-018-418-4
7	1165	28.8	798	7	US-10-018-418-2
8	1148.5	28.4	799	6	US-09-952-677-6
9	1142.5	28.3	690	7	US-10-044-543-6
10	1120.5	27.7	597	7	US-10-018-418-6
11	875.5	21.6	616	6	US-10-044-543-14
12	863	21.3	564	6	US-09-595-328C-275
13	863	21.3	611	6	US-09-595-328C-274
14	862.5	21.3	528	6	US-09-595-328C-276
15	860	21.3	614	7	US-10-044-543-18
16	852.5	21.1	600	7	US-10-044-543-22
17	828.5	20.5	615	7	US-10-044-543-2
18	590.5	14.6	961	6	US-09-708-427-12954
19	590.5	14.6	1027	6	US-09-708-427-12953
20	590.5	14.6	1071	6	US-09-708-427-12952
21	568	14.0	533	7	US-10-044-543-10
22	535	13.2	867	6	US-09-708-427-7513
23	535	13.2	900	6	US-09-708-427-7512
24	535	13.2	1025	6	US-09-708-427-7511
25	496.5	12.3	1059	7	US-10-018-418-10

26 495.5 12.3 1628 7 US-10-018-418-8  
27 419 10.4 474 7 US-10-007-693-99  
28 389.5 9.6 476 7 US-10-007-693-69  
29 138 3.4 33 6 US-09-118-990A-29  
30 132.5 3.3 443 6 US-09-708-427-50399  
31 131.5 3.3 733 6 US-09-708-427-12776  
32 131.5 3.3 756 6 US-09-708-427-12775  
33 131.5 3.3 796 6 US-09-708-427-10650  
34 130.5 3.2 2110 6 US-09-614-150-1023  
35 129.5 3.2 515 6 US-09-620-394B-6751  
36 129.5 3.2 523 6 US-09-620-394B-6750  
37 129.5 3.2 525 6 US-09-620-394B-6749  
38 127 3.1 632 6 US-09-708-427-12777  
39 127 3.1 695 6 US-09-708-427-10651  
40 125.5 3.1 732 6 US-09-708-427-11216  
41 125.5 3.1 746 6 US-09-708-427-11215  
42 123 3.0 27 7 US-10-018-418-53  
43 123 3.0 358 6 US-09-897-516-7504  
44 122 3.0 191 6 US-09-924-358-22  
45 120.5 3.0 194 6 US-09-708-427-56444

## ALIGNMENTS

## RESULT 1

US-09-674-824-2

; Sequence 2, Application US/09674824

; GENERAL INFORMATION:

; APPLICANT: Lorz, et al

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH CODE FOR ENZYMES DERIVED FROM WHE

; FILE REFERENCE: 1998/M 205/ FLH514413-3848

; CURRENT APPLICATION NUMBER: US/09/674,824

; CURRENT FILING DATE: 2000-02-16

; PRIOR FILING DATE: 1999-05-07

; PRIOR FILING DATE: 1998-05-07

; PRIOR FILING DATE: 1998-05-08

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 2

; LENGTH: 756

; TYPE: PRT

; ORGANISM: Triticum aestivum

US-09-674-824-2

Query Match 100.0%; Score 4044; DB 5; Length 756;  
Best Local Similarity 100.0%; Pred. No. 7.7e-317; Indels 0; Gaps 0;  
Matches 756; Conservative 0; Mismatches 0;

Qy 1 MAATGVGAGCLAPSVRLRADPATAARASACVVRARLRRLARGRYVAELSRGPAARPAQ 60  
Db 1 MAATGVGAGCLAPSVRLRADPATAARASACVVRARLRRLARGRYVAELSRGPAARPAQ 60  
Qy 61 QOLAPPLVPGFLAPPAPAPASPTOPPLPDAGVGLAPDLLLEGTAEDSIDSIIVAAAS 120  
Db 61 QOLAPPLVPGFLAPPAPAPASPTOPPLPDAGVGLAPDLLLEGTAEDSIDSIIVAAAS 120  
Qy 121 EODSEIMDANEQOKVTRTSIVFTVGEAAPYAKSGGLGDCVCSLPALAAARGHRVWVMP 180  
Db 121 EODSEIMDANEQOKVTRTSIVFTVGEAAPYAKSGGLGDCVCSLPALAAARGHRVWVMP 180  
Qy 181 RYLNGSSDKNYAKALYIAKHKIKPCFGSGHEVTFEYRDNDVWVFDHPSYHRPGSLYG 240  
Db 181 RYLNGSSDKNYAKALYIAKHKIKPCFGSGHEVTFEYRDNDVWVFDHPSYHRPGSLYG 240  
Qy 241 DNFCAFQDNQFRYTLCLYACAEAPLILELGGYIYQNCMFVNDWHDHSLVPLVLAAYRP 300  
Db 241 DNFCAFQDNQFRYTLCLYACAEAPLILELGGYIYQNCMFVNDWHDHSLVPLVLAAYRP 300  
Qy 301 YGVTRDSRSTLVIIHNLAHQGVPEASTYDGLPPEWYGALEWFPPEWARRHALDKGEAVN 360



|||||  
Db 301 YGYRDSRSLVTHLAHQGVESPASTYDPLGLPPEWYGALEWVPEWARRHALDKCEAVN 360  
QY 361 FLKGAVVADRIYVTSOGYSWEVTTAEGGGLNELLSRKSVLNGVINGIDINDWNPSTD 420  
Db 361 FLKGAVVADRIYVTSOGYSWEVTTAEGGGLNELLSRKSVLNGVINGIDINDWNPSTD 420  
QY 421 KCLPHYSVDLSGKAKKAELOKELGVPVREDVPLIGFIRLDYQKGDILIKMAIPELM 480  
Db 421 KCLPHYSVDLSGKAKKAELOKELGVPVREDVPLIGFIRLDYQKGDILIKMAIPELM 480  
QY 481 REDVQFVMLGSGDPIFGWMRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFEP 540  
Db 481 REDVQFVMLGSGDPIFGWMRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFEP 540  
QY 541 CGLNQLYAMQYGTVPVYVHGTGGLRDVETNPFCAGKEEGTGHAFSPLTVDKMLWALRTA 600  
Db 541 CGLNQLYAMQYGTVPVYVHGTGGLRDVETNPFCAGKEEGTGHAFSPLTVDKMLWALRTA 600  
QY 601 MSTPREHKPSWGLMKRGMTKDHTWDHAPSTSRSSGSPSWTNPSTCRRGLGRSKCESPS 660  
Db 601 MSTPREHKPSWGLMKRGMTKDHTWDHAPSTSRSSGSPSWTNPSTCRRGLGRSKCESPS 660  
QY 661 ALKTSSSFRGPEGYPTLCRCPATVESQACLLWFAGRSRYDGCAAAAYTASGGRQLQFW 720  
Db 661 ALKTSSSFRGPEGYPTLCRCPATVESQACLLWFAGRSRYDGCAAAAYTASGGRQLQFW 720  
QY 721 GIRKCAAGWLTAKHSDGSLVSURVTAIRNQLVTL 756  
Db 721 GIRKCAAGWLTAKHSDGSLVSURVTAIRNQLVTL 756

RESULT 2

US-09-952-677-2  
; Sequence 2, Application US/09952677  
; GENERAL INFORMATION:  
; APPLICANT: Block, Martina  
; Lorz, Horst  
; Luticke, Stephanie  
; Walter, Lennart  
; Froberg, Claus  
; Kosmann, Jens  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES  
; FROM WHEAT WHICH ARE INVOLVED IN STARCH  
; SYNTHESIS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James F. Haley, Jr., C/o Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/952,677  
; FILING DATE: 14-Sep-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/196,390  
; FILING DATE: 19-Nov-1998  
; APPLICATION NUMBER: DE 196 21 588.9  
; FILING DATE: 29-MAY-1996  
; APPLICATION NUMBER: DE 196 36 917.7  
; FILING DATE: 11-SEP-1996  
; APPLICATION NUMBER: PCT/EP97/02793  
; FILING DATE: 28-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley, Jr., James F.  
; REGISTRATION NUMBER: 27,794

; REFERENCE/DOCKET NUMBER: AGREVO-9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 596-9000  
; TELEFAX: (212) 596-9090  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 671 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-952-677-2

Query Match 89.3%; Score 3612; DB 6; Length 671;  
Best Local Similarity 100.0%; Pred. No. 3.7e-282;  
Matches 671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 86 TQPLPDAGVGLAPDLLLEGTAEDSIDSIIVAASEQDSEINDANEOPQAKVTRISIVFT 145  
Db 1 TQPLPDAGVGLAPDLLLEGTAEDSIDSIIVAASEQDSEINDANEOPQAKVTRISIVFT 60  
QY 146 GEAPYAKSGGLDVCGLPIALAAARHVRVMPRYLNGSSDKNYAKALYAKHIKIP 205  
Db 61 GEAPYAKSGGLDVCGLPIALAAARHVRVMPRYLNGSSDKNYAKALYAKHIKIP 120  
QY 206 FGSHEVTFEYHRNDVWVVDHPHPSYHRPGSLYGNFAGFGNQFRYTLACYACEAPL 265  
Db 121 FGSHEVTFEYHRNDVWVVDHPHPSYHRPGSLYGNFAGFGNQFRYTLACYACEAPL 180  
QY 266 ILEGGYIYQNCMFVYNDWVHSLVPLVLLAAKYRPGYVYRDSRSTLVINLHAHQVEPAS 325  
Db 181 ILEGGYIYQNCMFVYNDWVHSLVPLVLLAAKYRPGYVYRDSRSTLVINLHAHQVEPAS 240  
QY 326 TYPDLGLPPEWYGALEWVPEWARRHALDKGEAVNFKGAVVTADRIYVTSQYSWEVTT 385  
Db 241 TYPDLGLPPEWYGALEWVPEWARRHALDKGEAVNFKGAVVTADRIYVTSQYSWEVTT 300  
QY 386 AEGQGLNELLSRKSVLNGVINGIDINDWNPSTDKLPHYSVDLSGKAKKAELOKE 445  
Db 301 AEGQGLNELLSRKSVLNGVINGIDINDWNPSTDKLPHYSVDLSGKAKKAELOKE 360  
QY 446 LGLPVRDVPVLIQFIRLDYQKGDILIKMAIPELMREDVQFVMLGSGDPIFGWMRSTES 505  
Db 361 LGLPVRDVPVLIQFIRLDYQKGDILIKMAIPELMREDVQFVMLGSGDPIFGWMRSTES 420  
QY 506 SYKDKFRGWGFSVPVSHRITAGCDILLMPSPRPEPGLNQLYAMQYGTVPVYVHGTGGLRD 565  
Db 421 SYKDKFRGWGFSVPVSHRITAGCDILLMPSPRPEPGLNQLYAMQYGTVPVYVHGTGGLRD 480  
QY 566 TVETFPFGAKGEEGTGNAFSPVTVDKMLWALRTAMSTFRHKPSWGLMKRGMTKDHTW 625  
Db 481 TVETFPFGAKGEEGTGNAFSPVTVDKMLWALRTAMSTFRHKPSWGLMKRGMTKDHTW 540  
QY 626 DHAPSSTSRSSGSPSWTNPSTCRRGLGRSKCESPSALKTSSSFRGPEGYPTLCRCPATV 685  
Db 541 DHAPSSTSRSSGSPSWTNPSTCRRGLGRSKCESPSALKTSSSFRGPEGYPTLCRCPATV 600  
QY 686 ESQCACLLWFAGRSRYDGCAAAAYTASGGRQLQFWGTRKCAAGWLTAKIHSDGSLSVRV 745  
Db 601 ESQCACLLWFAGRSRYDGCAAAAYTASGGRQLQFWGTRKCAAGWLTAKIHSDGSLSVRV 660  
QY 746 TAEIRNQLVTL 756  
Db 661 TAEIRNQLVTL 671

RESULT 3  
US-09-508-377-14  
; Sequence 14, Application US/09508377  
; GENERAL INFORMATION:  
; APPLICANT: KALEEN, ZHONGYILI  
; APPLICANT: MORELL, MATTHEW



APPLICANT: RAHMAN, SADEOUR  
TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS  
FILE REFERENCE: 054270/0126  
CURRENT APPLICATION NUMBER: US/09/508,377  
CURRENT FILING DATE: 2000-06-09  
PRIOR APPLICATION NUMBER: AU PP 2509  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: PCT/AU98/00743  
PRIOR FILING DATE: 1998-09-11  
PRIOR APPLICATION NUMBER: AU PP 9108  
PRIOR FILING DATE: 1997-09-12  
NUMBER OF SEQ ID NOS: 71  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 14  
LENGTH: 647  
TYPE: PRT  
ORGANISM: Triticum tauschii  
US-09-508-377-14

Query Match 82.8%; Score 3348; DB 6; Length 647;  
Best Local Similarity 99.7%; Pred. No. 6e-261;  
Matches 626; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAATGVGAGCLAPSVLRADPATAARASACVVRARLRRLARGRYVAELSRGPAARPAQ 60  
Db 1 MAATGVGAGCLAPSVLRADPATAARASACVVRARLRRLARGRYVAELSRGPAARPAQ 60

QY 61 QOLAPPLVPGFLAPPPAPAPQAPTOPPLPDAGVGLAPDLLLEGIAEDSIDSIIVAS 120  
Db 61 QOLAPPLVPGFLAPPPAPAPQAPTOPPLPDAGVGLAPDLLLEGIAEDSIDSIIVAS 120

QY 121 EQSEIMDANEQOAKVTRISIVFTGEAAPYAKSGGLDVCGLPITALAARGHRVMVMP 180  
Db 121 EQSEIMDANEQOAKVTRISIVFTGEAAPYAKSGGLDVCGLPITALAARGHRVMVMP 180

QY 181 RYLNGSSDKNYAKALYAKHIKIPCGGSHVETFFHEYRDNDVMDVDPHPSYHRPGSLYG 240  
Db 181 RYLNGSSDKNYAKALYAKHIKIPCGGSHVETFFHEYRDNDVMDVDPHPSYHRPGSLYG 240

QY 241 DNFGAGDNQFRYTLACYACEAPLLELGGYIYGONCMFVNDWHSALVPLVLAARYP 300  
Db 241 DNFGAGDNQFRYTLACYACEAPLLELGGYIYGONCMFVNDWHSALVPLVLAARYP 300

QY 301 YGYRDSRSTLVTHNLAHQGVPEASTYPDLGLPPEWYGALEWVPEWARRHALDKGEAVN 360  
Db 301 YGYRDSRSTLVTHNLAHQGVPEASTYPDLGLPPEWYGALEWVPEWARRHALDKGEAVN 360

QY 361 FLKGAVVTADRIVTVSQGYSWEVTTAEGGQGLNELLSRKSVLNGTVNGIDINDNPTTD 420  
Db 361 FLKGAVVTADRIVTVSQGYSWEVTTAEGGQGLNELLSRKSVLNGTVNGIDINDNPTTD 420

QY 421 KCLPHYSYDDLSGKAKCAEQLKELGLPVRDVPVLIGFGRDLYOKGIDLIKMAIPELM 480  
Db 421 KCLPHYSYDDLSGKAKCAEQLKELGLPVRDVPVLIGFGRDLYOKGIDLIKMAIPELM 480

QY 481 REDVQVLMGSDPIFEGHMRSTESSYKDKFRGWGFSVPVSHRTAGCDILLMPSRREP 540  
Db 481 REDVQVLMGSDPIFEGHMRSTESSYKDKFRGWGFSVPVSHRTAGCDILLMPSRREP 540

QY 541 CGLNQLYAMQYGTVPVHGTGGLRDTVETFPFGAKGEGTGWAFSPLTVDKMLWALRTA 600  
Db 541 CGLNQLYAMQYGTVPVHGTGGLRDTVETFPFGAKGEGTGWAFSPLTVDKMLWALRTA 600

QY 601 MSTFREHKPSWGLMKRGMTKDHTWDHA 628  
Db 601 MSTFREHKPSWGLMKRGMTKDHTWDHA 628

## RESULT 4

US-09-931-297-2  
Sequence 2, Application US/09931297  
GENERAL INFORMATION:

APPLICANT: Jens Kossmann  
Claus Froberg  
TITLE OF INVENTION: Nucleic acid molecules encoding soluble starch synthases from maize  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: James F. Haley, Jr., c/o FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)  
CURRENT APPLICATION DATA: US/09/931,297  
APPLICATION NUMBER: 09/192,909  
FILING DATE: <Unknown>  
APPLICATION NUMBER: DE 196 19 918.2  
FILING DATE: 17-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley, Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: GFB-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000  
TELEFAX: 212-96-9090  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 649 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-931-297-2

Query Match 68.0%; Score 2749; DB 6; Length 649;  
Best Local Similarity 82.5%; Pred. No. 9.4e-213;  
Matches 518; Conservative 34; Mismatches 62; Indels 14; Gaps 4;

QY 4 TGVGAGCLAPSVLRADPATAARASACVVRARLRRLARGRYVAELSRGPAARPAQ 60  
Db 14 SAVGAACLL-----LARAANPAAVGDRAARPRRLQVRRLRCVAELSRGPAARMP 65

QY 61 QOLAPPLVPGFLAPPPAPAPQAPTOPPLPDAGVGLAPDLLLEGIAEDSIDSIIVAS 120  
Db 66 ALLAPPLVPGFLA-PPAETGEPALTPPPVPDAGLVGLVE--PEGIAEGSIDNTVVVAS 122

QY 121 EQSEIMDANEQOAKVTRISIVFTGEAAPYAKSGGLDVCGLPITALAARGHRVMVMP 180  
Db 123 EQSEIVVVKEQARAKVTQNVFTGEASPYAKSGGLDVCGLPITALAARGHRVMVMP 182

QY 181 RYLNGSSDKNYAKALYAKHIKIPCGGSHVETFFHEYRDNDVMDVDPHPSYHRPGSLYG 240  
Db 183 RYLNGTSDKNYANAFYTEKHIRIPCFGGEHEVTFHEYRDSVDVMDVDPHPSYHRPGSLYG 242

QY 241 DNFGAGDNQFRYTLACYACEAPLLELGGYIYGONCMFVNDWHSALVPLVLAARYP 300  
Db 243 DKFGAGDNQFRYTLACYACEAPLLELGGYIYGONCMFVNDWHSALVPLVLAARYP 302

QY 301 YGYRDSRSTLVTHNLAHQGVPEASTYPDLGLPPEWYGALEWVPEWARRHALDKGEAVN 360  
Db 303 YGYKDSRSILVTHNLAHQGVPEASTYPDLGLPPEWYGALEWVPEWARRHALDKGEAVN 362

QY 361 FLKGAVVTADRIVTVSQGYSWEVTTAEGGQGLNELLSRKSVLNGTVNGIDINDNPTTD 420  
Db 361 FLKGAVVTADRIVTVSQGYSWEVTTAEGGQGLNELLSRKSVLNGTVNGIDINDNPTTD 420



Db 363 FLKAVVYADRIIVTSKGYSWETTAAEGGGLNELLSSRKSVLNGIVNGIDINDWNPATD 422  
Qy 421 KCLPHYSVDDLSGKAKKAELOKELGVLVREDVPLIGTGRDYOKGIDLIKMAIPELM 480  
Db 423 KCIPCHYSVDDLSGKAKKAELOKELGVLVREDVPLIGTGRDYOKGIDLIKMAIPELM 482  
Qy 481 REDVQVFMGSGDPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFP 540  
Db 483 REDVQVFMGSGDPELEDWNRSTESIFKDKFRGWGFSVPVSHRITAGCDILLMPSRFP 542  
Qy 541 CGLNQLYAMQYGVVPHVHTGGLRDVTETFPNFGAKGEGTGWAFSLVVDKMLWALRTA 600  
Db 543 CGLNQLYAMQYGVVPHVHTGGLRDVTETFPNFGAKGEGTGWAFSLVVDKMLWALRTA 602  
Qy 601 MSTFREHKPSWEGMLKRGMTKDHWTWDA 628  
Db 603 ISTYREHKSSWEGMLKRGMTKDHWTWDA 630  
RESULT 5  
US-10-044-543-26  
; Sequence 26, Application US/10044543  
; GENERAL INFORMATION:  
; APPLICANT: Singletary, George  
; APPLICANT: Zhou, Lan  
; TITLE OF INVENTION: Novel Starch Synthese Polynucleotides  
; FILE REFERENCE: 1144D  
; CURRENT APPLICATION NUMBER: US/10/044, 543  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: 09/388,743  
; PRIOR FILING DATE: 1999-09-02  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 26  
; LENGTH: 801  
; TYPE: PRT  
; ORGANISM: Typha latifolia  
US-10-044-543-26

Query Match 29.6%, Score 1196.5; DB 7; Length 801;  
Best Local Similarity 42.6%; Pred. No. 1e-87;  
Matches 269; Conservative 80; Mismatches 200; Indels 83; Gaps 15;  
Qy 42 GRYVAELREGPAARPAQQQLAP-----PLVPGFLAPPPP-----APAQSPAPT 86  
Db 109 GRYKS-----GAVPNYSOLAQDDTSENPLVNSFGSPKDNVEAFVQYRQSAVDAG 161  
Qy 87 QPPLPDAGVGLAPDLLLEG-----IAEDSDTSIIVAASEODSEIMDANQOP- 133  
Db 162 RPEESLGTTKLSPFYLEAESDGAKEAEDLVEAKLDSVHV---KDDLNPGEENEVPL 218  
Qy 134 --QAKVTSISIVFTGEAAPAKSGGLGVCCLPTALAARGHRVVMVPRYLNGSSDKNY 191  
Db 219 PLAGANVMNIIVAAECAPSKTGLGVDVAGALPKALARRHRVVMVAPRGNTAEPODI 278  
Qy 192 AKALYTAKHKIPCFGSGSHSVTFEYRDNDVMDVFDVHPSY-HRPGSLYGNFGAFGDNQ 250  
Db 279 GVRKYKVVH-----GQDMEVTFYHAYIDGVDFVMDSPDFRHRGNRIYECN---RVDIL 329  
Qy 251 FRYTLCLYAAECAPLILELGGYIYGQ--NCMFVNDWNASHLVPLLAARYPGYVYRDSRS 309  
Db 330 KRMILFCKAAAEVPMVHVPFCGGCYGDGNLAFITNDWHTALLPVYLKAYIRONGLMKYARS 389  
Qy 310 TLVIHNLAHQGVPERASTYDPLGLPPEWYGALEWVPFPAWRHALD-----KGEAYNF 361  
Db 390 VLVIHNLAHQGVPERASTYDPLGLPPEWYGALEWVPFPAWRHALD-----HYLDLFRYPVGEHLNI 433  
Qy 362 LKAVVYADRIIVTSKGYSWETTAAEGGGLNELLSSRKSVLNGIVNGIDINDWNPATD 421  
Db 434 FAAGLKTADRVVTVSHGYAWELKTSEGGWGLHEIINESNWKFGQIVNGIDAKESWSPEDV 493

Qy 422 CLUPH-----HYSVDDLS-CKAKKAELOKELGVLVREDVPLIGTGRDYOKGIDLIKMAI 476  
Db 494 HLKSDGYTYNSLDTLEMGKPVCKAALQREVGLPVHDNVPILIAFICRLDQHGKVDLIAEAM 553  
Qy 477 PELMREDVQVFMGSGDPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPS 536  
Db 554 PWIVSHDVQVFMGSGDPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPS 613  
Qy 537 REPCGLNQLYAMQYGVVPHVHTGGLRDVTETFPNFGAKGEGTGWAFSLVVDKMLWA 596  
Db 614 REPCGLNQLYAMQYGVVPHVHTGGLRDVTETFPNFGAKGEGTGWAFSLVVDKMLWA 670  
Qy 597 LRTANSTFREHKPSWEGMLKRGMTKDHWTWDA 628  
Db 671 LNNCLNTYWNKYDSWKGLQTRGMQDLSMDNA 702  
RESULT 6  
US-10-018-418-4  
; Sequence 4, Application US/10018418  
; GENERAL INFORMATION:  
; APPLICANT: Morell, Matthew  
; APPLICANT: Li, Zhongyi  
; APPLICANT: Rahman, Sadequr  
; APPLICANT: Appels, Rudolph  
; TITLE OF INVENTION: Genes Encoding Wheat Starch Synthases and Uses Therefor  
; FILE REFERENCE: 127-01  
; CURRENT APPLICATION NUMBER: US/10/018,418  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: PCT AU00/00385  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: AU PQ0052/99  
; PRIOR FILING DATE: 1999-04-29  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 799  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-10-018-418-4

Query Match 29.0%, Score 1174.5; DB 7; Length 799;  
Best Local Similarity 37.1%; Pred. No. 5.9e-86;  
Matches 286; Conservative 91; Mismatches 223; Indels 171; Gaps 19;  
Qy 7 GAGCLA-----PSVRLRADPATAARA-----SACVVRRLRLRLRGRVVAE---- 47  
Db 36 GAGRLHWPWPQPPTARDGCVAAAGKKDARVDDDAASARQPRARRGGAATKVAERRDP 95  
Qy 48 ---LSRE-----GPAARPAQQQLAPPLVPGFLAPP-----PAPAO 82  
Db 96 VKTLDRDAEAGGAPAPAPAPQDAARP--PSMNGTPVNGENKSTGGGATKDSGLPAPARA 153  
Qy 83 PAP--TQPLPDAGVGL-----APDLLLEGAEDSIDSIIVAAEQDSEI----- 126  
Db 154 PHESTQNRVNGENKANVASPPTSIAEVVAPDSAAATISIDKAPESVVPAAEKPPSSGS 213  
Qy 127 -----MDANEQPOAKV----- 137  
Db 214 NFVVSASAPRLIDSDVEPELKKGAVIVEAPNPKALSPAPPAVQBDLWDFKYYIGFEE 273  
Qy 138 -----TRSVFVFTGEAAPYAKSGGLGVDVCGSL 164  
Db 274 PVEAKDDGWAADDAGSEFHHQNHDSGLAGENVMNVVVVAAECSPWCKTGGGLDVAGAL 333  
Qy 165 PIALAARGHRVVMVPRYLNGSSDKNYAKALYTAKHKIPCFGSGHSHVTFEYRDNDV 224  
Db 334 PKALAKRGHRVVMVPRY----GDYEEAYDVGVRKYYK--AAGQDMEVNTFYHAYIDGVDF 387  
Qy 225 VFVDHPSY-HRPGSLYGNFGAFGDNQFRYTLCLYAAECAPLILELGGYIYGQ--NCMFVV 282  
Db 388 VFIDAPLFRHQRQEDIYG----GSRQEIIMKRMILFCKAAAEVPMVHVPFCGGVPGVGDGNLFIA 444



```

274 274  VEAKDGRAVDADGAFSEFHQNDHSGPLAGENVMNVVVAAECSPWCKTGGLGDVACALP 333
QY 166  IALAARGHRVMMVPRLNGSSDKNYAKALYAKHIKIPCEGSGHEVTFPHETRDNDVWV 225
Db 334  KALAKRHRVMMVVPRIY-----GDYEADYDVGVRKYK--AAGQDMENVNYPHAYIDGVDF 387
QY 226  FVDHPSY-HRPGSLYGNFAGFDNQFRYTLILCYAAECAPLILELGLGYIYGQ-NCMPFVN 263
Db 388  FIDAPLFRHQEDIYG--GSRQEIIMKRMILFCKAAVEVPHVPCGGVPYGDGNLVIAN 444
QY 284  DWHASLPVLLAAKRYGYVYRDSRSTLVTHNLAHOGVEPASTYVDLGLPPEWYGALEWV 343
Db 445  DWHALLPVLKAYIRDHGLMQYTRSTWVJHNTAHQGRGPDVEFTELPHEHYL-----498
QY 344  FPEWARHALDKGEAVNFKGAVVTADRIVTVSQGYSWEVTTAEGOGNELLSRKSXYL 403
Db 499  --EHFRLYDPVGGEHANYFAAGLKMAOVVVVSPGYLWELKTVEGGWGLHDIIIRQNDWT 556
QY 404  NGIVNGIDINDWPTTDKCLPH-----HYSVDDL-SGAKACKAELQELGLPVRVEDVPLIG 458
Db 557  RGIVNGIDINMEWNPVEDVHLKSDGYTNFSLGTLDSGKCKEALQRELGLQVRGDPVLLG 616
QY 459  FIGRLDYQKIDILKMAIPELMREDOYVMLSGSDPIFGGWRSTESSYKDKFRGWGFS 518
Db 617  FIGRLDQKQVEIADAMPWISQDVQLVNLGTGRHDLGMLRHFEREHHDKVRGWGFS 676
QY 519  VPVSHRITAGCDILLMPSRFPCCGLNOLYAMQYGVVYVHGTGGLRDTVETENPFCAKGE 578
Db 677  VRLAHRITAGADALLMPSRFPCCGLNOLYAMAYGTVPVYHVGGLRDTVPPFDPE--NH 733
QY 579  EGTGWAFSPLTVDKMLWALRTAMSTREHKPSWEGMLKRGMTKDHWTWHA 628
Db 734  SGLGWTFDRAEQKLIEALCHCLRTYDYKESWNRGLQERMSQDSWEHA 783

RESULT 8
US-09-952-677-6
; Sequence 6, Application US/09952677
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; Lorz, Horst
; Lutticke, Stephanie
; Walter, Lennart
; Frohberg, Claus
; Kossmann, Jens
;
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; FROM WHEAT WHICH ARE INVOLVED IN STARCH
; SYNTHESIS
;
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/952,677
; FILING DATE: 14-Sep-2001
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,390
; FILING DATE: 19-Nov-1998
; APPLICATION NUMBER: DE 196 21 588.9
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: DE 196 36 917.7
; FILING DATE: 11-SEP-1996
; APPLICATION NUMBER: PC/EP97/02793
; FILING DATE: 11-SEP-1996

```



ATTORNEY/AGENT INFORMATION:  
NAME: Haley, Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: AGREVO-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 799 amino acids  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-952-677-6

Query Match 28.4%; Score 1148.5; DB 6; Length 799;  
Best Local Similarity 37.0%; Pred. No. 7.2e-84;  
Matches 286; Conservative 86; Mismatches 226; Indels 175; Gaps 20;

Qy 7 GAGCLA-----PSVRLRADPATAA-----RASACVVRARLRRLARGRYVAE----- 47  
Db 36 GAGRLHWPWPPTARDGVAALAAAGKKDAGIDDAASVQPRALRGGAATKVAERRDP 95  
Qy 48 ---LSRE-----GAARPAQOQQLAPLPVPGFLAPPP-----PAPA 80  
Db 96 VKTLDRDAEAGGSPPAARDQDAARPPSMNGM-----PVNGENKSTGGGATKDSGLPTPA 151  
Qy 81 QSPAP-TQPLPDAGVGL-----APDLLLEGIAEDSIDSIIV----- 117  
Db 152 RAPHESTQNRAPVNGENKANTASPTSTIAEAAASDAATISISOKAPESVVPABKTPSS 211  
Qy 118 -----AASEQDSEIMDANEQ-----POAKV----- 137  
Db 212 GSNFESSASAGSDTSDVEQELKKGAVVEEAPKRALSPAPAPVQEDLWFKKYGIF 271  
Qy 138 -----TRISIVFTGEAAPYAKSGGLGVCG 162  
Db 272 BEPVENKODGRAVADAGSFEHHQNHDSGLAGENVNMVVAAPCSWCKTGGGLGVAG 331  
Qy 163 SLPIALARGHRVVMVPRYLNGSSDKNYAKALYAKHIKPCGGSHSEVTFEHEYRDNV 222  
Db 332 ALPKALAKRGHRVMVWPY-----GDYBEAVDVGVRKYK---AGQDMENVFIHAYIDGV 385  
Qy 223 DMVFDVHPSY-HRPGSLVGDNPGAFGDNQFRYTLACYAAEAPLLELGGYIYGO-NCMF 280  
Db 386 DFVFDAPLFRHQEDYIY-----GSQEIIMKRMILFCKAAVEVPWHVPCGVPYGDGNLVF 442  
Qy 281 VVNDHASLVPLLAAKYPYGVYRDSRSTLVIHNLAHQGVPEPASTYFDLGLPPEWYCAL 340  
Db 443 IANDWHTALLPVLYKAYYRDHGLMOTYRSMVIMHIAHOGRPVDFEFPTELPEHYL--- 499  
Qy 341 EWVPEWARHRLDKGEAVNFKLGAVVTADRTVTSQYSWEVTTAEGGQGLNELLSRK 400  
Db 500 -----EHFRIDYDPVGEHANFYAAGLKMDQVYVSPGLWELKTVEGWLHDIIRQND 554  
Qy 401 SVLNGVINGIDINDNPTDKLPH-----HYSVDLL-SCKAKCKAELQELGLPVREDVP 455  
Db 555 WKTRGVINGIDNMEWNPEDVAHLKSDGYTNFSLRTLDSGKQCKEALQELGLQVRADVP 614  
Qy 456 LIGFIRGLDYQKIDILKKAIPELMREDVQFVNLGSGSDPIFEQWNRSTESSYKDKFRGWV 515  
Db 615 LLGFIRGLDGOKGVEIIADAMPVIVSQDQLVNLGTGRHDLSEMLQGFEREHHDKVRGWV 674  
Qy 516 GFSVPVSHRITAGCDILLMPSPFCGLNQLYAMQYGTVPVHVHGTGLDRDTVETNPFKA 575  
Db 675 GFSVRLAHRITAGADALLMPSPFCGLNQLYAMAYGTVPVHVHGTGLDRDTVPPDFPF--- 732  
Qy 576 KGEETGWAFSPITVDKMLWALTAMSTFREHKPSWEGLMKRGMTKDHWTWDA 628  
Db 733 -NHSGLWTDFRAEAHKLIEALGHCLRTYRDFKESWRALOERMSQDFSWEHA 784

RESULT 9

US-10-044-543-6  
; Sequence 6, Application US/10044543  
; GENERAL INFORMATION:  
; APPLICANT: Singletary, George  
; APPLICANT: Zhou, Lan  
; TITLE OF INVENTION: Novel Starch Synthase Polynucleotides  
; TITLE OF INVENTION: and Their Use in the Production of New Starches  
; FILE REFERENCE: 1144D  
; CURRENT APPLICATION NUMBER: US/10/044,543  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: 09/388,743  
; PRIOR FILING DATE: 1999-09-02  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Curcuma zedoaria  
US-10-044-543-6

Query Match 28.3%; Score 1142.5; DB 7; Length 690;  
Best Local Similarity 47.5%; Pred. No. 1.8e-83;  
Matches 238; Conservative 76; Mismatches 150; Indels 37; Gaps 11;

Qy 140 STVFVTGEAAPYAKSGGLGVCGSLPIALAAAGHRVVMVPRYLNGSSDKNY--AKALYT 197  
Db 200 NIILVAEAPKSPKSGGLGVCGALPKALAKRGHRVMVSPRY-----GNYPEKPEIGN 253  
Qy 198 AKHIKPCGGSHSEVTFEHEYRDNVDMVFDHPSYHRPGS-LYGDNFGAFGQNFRTYLL 256  
Db 254 LKRYKVD--GQDMEIKYHYTIDSVDFIDSPFRHIGNDIYGN---RVDILKRWVLF 308  
Qy 257 CYAAEAPLLELGGYIYGO-NCMFVYVNDWHSVLPVLLAAKYPYGVYRDSRSTLVIH 315  
Db 309 CKAAVEVPWHVPCGFCYGDGNLVFIANDWHTSLLPVYLKACFRDRGLMTYARCLLVIH 368  
Qy 316 LAHQGVPEASTYFDLGLPPEWYCALEWPEWARHRLDK---GEAVNFKLGAVVTADRI 372  
Db 369 IAHQGRGLPDDFSYDLPDHDHSFR-----LDDPVGGEHFNIFAAGIRANDRV 417  
Qy 373 VTSQYSWEVTTAEGGQGLNELLSRKSVLNGVINGIDINDNPTDKLPH-----HVS 428  
Db 418 VTSYHGYAWELKTSEGQWGLHEIINECHWKHFHGLVINGIDTHSNPKFADHLSNDSGYTNET 477  
Qy 429 VDDL-SCKAKCKAELQELGLPVREDVPLIGFTIGRLDYQKIDILKKAIPELMREDVQFV 487  
Db 478 LETLEMGKAOCKAALQREFGLPVRDDVPILAFIQLRDLHQKIDILAEAMHMLVGGDLQII 537  
Qy 488 MLGSGDPIFEQWNRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSPFCGLNQLY 547  
Db 538 MLGTGRDLEMLRRFREREHKGKVRGWGFSVKMAHRTAGADALLMPSPFCGLNQLH 597  
Qy 548 AMQYGTVPVHVHGTGLDRDTVETNPFKAGEETGWAFSPITVDKMLWALTAMSTFREH 607  
Db 598 AMYGTIPVHVHAGGLRDTVQQDFPF---NETGLGWTDFRAEAHRMTVALGHCLNTYRNY 654  
Qy 608 KPSWEGLMKRGMTKDHWTWDA 628  
Db 655 KESWGLQKRGMMQDLSWESA 675

RESULT 10

US-10-018-418-6  
; Sequence 6, Application US/10018418  
; GENERAL INFORMATION:  
; APPLICANT: Morell, Matthew  
; APPLICANT: Li, Zhongyi  
; APPLICANT: Rahman, Sadequr  
; APPLICANT: Appels, Rudolph  
; TITLE OF INVENTION: Genes Encoding Wheat Starch Synthases and Uses Therefor



```

; FILE REFERENCE: 127-01
; CURRENT APPLICATION NUMBER: US/10/018,418
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: PCT AU00/00385
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: AU PQ0052/99
; PRIOR FILING DATE: 1999-04-29
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-018-418-6

```

```

Query Match      27.7%; Score 1120.5; DB 7; Length 597;
Best Local Similarity 41.7%; Pred. No. 8.6e-82;
Matches 254; Conservative 83; Mismatches 199; Indels 73; Gaps 16;

QY 37 RRLARGRYVAELSRGPAARPAQOQOOLAPPLVPG-----FLAPPPAPAPQAQSPAPT 86
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 30 QELKKAVVVE---EAP-----KPKALSPAPAPAVOEDLWDFKKYIGFEPVEAK----- 76
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 87 QPPLDAGVGEAPDLLLEGIAEDSDSIIVAASEQDSEIMDANEQOQAKVTRSIYFVVG 146
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 -----DDG-----RAVADD-----AGSPFHQNHDSGPLAGENV-MNVVVVAA 113
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 147 EAPYAKSGGLDVCGLSPALAAARGHRVVMVPRYLNGSSDKNYAKALYAKHIKIPCF 206
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 114 ECSPMCKTGGGLGVAGALPKALAKRGHRVVMVPRY-----GDYEEPTDGVRRYK--AA 167
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 207 GGSHEVTFEHEYRDNDVDFVDPHSY--HRPGSLYDGNFGAFGDNQFRYTLICVAAACEAPL 265
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 168 GQDMEVNYHAYIDGVDFVIDAPLFRHREDIYG---GSRQIMRMILFCKAAVEVPW 224
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 266 ILELGGYIYQ--NCMFVNDWHASLVPVLLAAKYRPGYVYRDSRSTLVIHNLAHQGVPA 324
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 225 HVPCGGVPYDGNLFTIANDWHTALLPVYLKAYYRDHGLMQYTRSIMVHNLIAHQGRGPV 284
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 325 STYPDGLPPEWGALEWFPPEWARHRLDKGEAVNLFKAGVYADRIYTVVSOGYSWEVT 394
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 285 DEFPFTPELPHYL-----EHFRLYDPVGGEHANYFAAGLKMADQVVVYVSPGYLWELK 336
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 385 TARGGGLNELSSKRSVLNGVINGIDINDWNPPTDKCLPH----HYSVDDL--SGRAKCK 439
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 337 TVEGGGLHDIIIRQNDWKTRGIVGIDNMENPEVDAHLKSDGYTNFSLRTLDSGRQCK 396
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 440 AELQKELGPLVREDVPLIGFIGRLDYQKIDILKMAIPELMREDVQFVMLGSGDPIFEGW 499
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 397 EALQRELGLQVRADVPLPLGFIGRLDQKQVEIADAMPVVSQDVQLVMLGTGRHDLISM 456
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 500 MRSTESSYKDKFGWGVFSVPVSHRITAGCDIILLMPSRPEPCGLNOLYAMQYGVTPVHG 559
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 457 LQHFEREHDKVRGWGVFSVRLAHRITAGADALLMPSRPEPCGLNOLYAMAYGVTPVVA 516
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 560 TGLRDTVFENFPKAGEGCTGWAESPLTVDKMLWALRTAMSTFREHKPSWEGLMKRG 619
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 517 VGLRDTVPFPDFP-----NHSLGLMTFDRAEAHLKLIIEALGHCLRTYRDFKESWALQERG 573
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 620 TKDHTWDHA 628
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 574 SQDFSWEHA 582
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 11
US-10-044-543-14
; Sequence 14, Application US/10044543
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: Novel Starch Synthesis Polynucleotides
; TITLE OF INVENTION: and Their Use in the Production of New Starches

```

```

; FILE REFERENCE: 1144D
; CURRENT APPLICATION NUMBER: US/10/044,543
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/388,743
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Canina edulis
US-10-044-543-14

Query Match      21.6%; Score 875.5; DB 7; Length 616;
Best Local Similarity 38.8%; Pred. No. 4.6e-62;
Matches 199; Conservative 90; Mismatches 171; Indels 53; Gaps 13;

QY 140 SIVFVTGEAAPYAKSGGLDVCGLSPALAAARGHRVVMVPRYLNGSSDKNYAKALYTAK 199
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 91 NLVFGVCEVAPWSKTGGGLDVLRLGLPPAMAAIGHRVMTVVPY-----DOYKDIWDTSV 144
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 200 HIKIPCFGSGSHEVTFPEHYRDNDVDFVDPHSY-----HRPGSLYGDNFGA--FGDNQPR 252
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 145 PVLEKVGDKIETVRFPHCYKRGVDRVDFVDPHMPLEKVMGKTGKLYGPVTGTDYADNQLR 204
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 253 YTLICVAAACEAPLILLEG--YI---YQNCMFVNDWHASLVPVLLAAKYRPGYVYRDS 307
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 205 FSLLCUAALEAPLLNLNNSKYSGPYGDDVFIANDWHASALLPCYLKTMYSQSHGIYMNA 264
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 308 RSTLVIHNLAHQGVPEASTYDPLGLPEWYGALEWYFPEWARHRLD---KGEAVNFKG 364
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 265 KVAFCIHNTAYOGRPAFSDPELLNLPNFKKSSFDEF-----DGYDKPVKGRKINMKA 317
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 365 AVVTADRIYTVVSOGYSWE--VTTAEGGQGLNELSSKRSVLNGVINGIDINDWNPPTDKCL 423
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 318 GIIECORCLTSPYYAQELVSGVEKVELGNIL--BMKTCIGVINGMDTTEWNPPLTDKYI 375
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 424 PHYSVDD--LSGKAKCAELQELGPLVREDVPLIGFIGRLDYQKIDILKMAIPELMRE 482
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 376 STNYDATTVDLAKPLCKEALQAEGLPVNKNKLVLAFAVGRLDQKSGDILAAAIPELCE 435
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 483 DVQFVMLGSGDPIFEGMRSSTESSYKDKFRGWGVFSVPVSHRITAGCDIILLMPSRPEPCG 542
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 436 NQQVILVGTGKKLESELTLLEEMFPDKFRAHLKFNVLAAHIMAGADILVIPSREPCG 495
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 543 LNOLYAMQYGVTPVVGHTGGLRDTVFENFPFGAKGEGCTGWAESPLTVD-----K 592
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 496 LIQLOAMRYGTLPMCSTTGGLVDTVK-----EGFTGFHMGPFPSVECDAYDKADVOK 546
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 593 MLWALRTAMSTFREHKPSWEGLMKRGMTKDHTW 625
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 547 IVETTRALKVY--GTPAFVEMIKMCMNODLSW 577
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 12
US-09-595-328C-275
; Sequence 275, Application US/09595328C
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-0946P
; CURRENT APPLICATION NUMBER: US/09/595,328C
; CURRENT FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3107
; SEQ ID NO 275
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature

```



LOCATION: 1..564  
OTHER INFORMATION: Xaa is any amino acid  
NAME/KEY: misc\_feature  
LOCATION: 1..564  
OTHER INFORMATION: Ceres Seq. ID 1009070  
US-09-595-328C-275

Query Match 21.3%; Score 863; DB 6; Length 564;  
Best Local Similarity 38.1%; Pred. No. 4.1e-61;  
Matches 203; Conservative 95; Mismatches 183; Indels 52; Gaps 15;  
QY 119 ASEQDSSEIMDANEOPQAKVTRTSIVFTGEEAAPYAKSGGLGVDGCSLPALAAARGHRVMV 178  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 123  
QY 179 MPRYLNGSSDKNAYAKALYTAHKIKIPCFGGSHEVTFEYRDNDVWVFDHPHY-----H 233  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 76  
QY 179 MPRYLNGSSDKNAYAKALYTAHKIKIPCFGGSHEVTFEYRDNDVWVFDHPHY-----H 233  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 76  
QY 234 RPSG-LYGDNFGA-FGDNQRYTLLCYAAACEAPLILELGG--YI---YGQNCMFVNDWH 286  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 130  
QY 234 RPSG-LYGDNFGA-FGDNQRYTLLCYAAACEAPLILELGG--YI---YGQNCMFVNDWH 286  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 130  
QY 287 ASLPVLLAAKYRPGYVYRDSRSTLVTHNLAHQGVPEASTYPDLGLPPEWYGALEWVFE 346  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 190  
QY 287 ASLPVLLAAKYRPGYVYRDSRSTLVTHNLAHQGVPEASTYPDLGLPPEWYGALEWVFE 346  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 190  
QY 347 WARRHALD-----KGEAVNFKAGVYADRTVTSOGYSWE-VTTAEGQGLNELSSR 399  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 246  
QY 347 WARRHALD-----KGEAVNFKAGVYADRTVTSOGYSWE-VTTAEGQGLNELSSR 399  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 246  
QY 400 KSVLNGIVNGIDINDNPTDKCLPHYSVDLLS-GKAKCKAEQLKELGLPVREDVPLIG 458  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 298  
QY 400 KSVLNGIVNGIDINDNPTDKCLPHYSVDLLS-GKAKCKAEQLKELGLPVREDVPLIG 458  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 298  
QY 459 FIGRLDYQKIDILKMAIPELMREDYQVFMGLSGDPFEGWMRSTESSYKDKFRGWGFS 518  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 358  
QY 459 FIGRLDYQKIDILKMAIPELMREDYQVFMGLSGDPFEGWMRSTESSYKDKFRGWGFS 518  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 358  
QY 519 VPVSHRTACDILLMPSRPEPCGLNOLYAMQYGTVPVHGTGGLRDTVET-----FNP 572  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 418  
QY 519 VPVSHRTACDILLMPSRPEPCGLNOLYAMQYGTVPVHGTGGLRDTVET-----FNP 572  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 418  
QY 573 FGAKGEGTGWASPLTVDMKMLALRTAMSTFREHKPSWEGLMKRGMTKDHWT 625  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 524

RESULT 13  
US-09-595-328C-274  
Sequence 274, Application US/09595328C  
GENERAL INFORMATION:  
APPLICANT: ALEXANDROV, Nikolai  
APPLICANT: BROVER, Vyacheslav  
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
FILE REFERENCE: 2750-0946P  
CURRENT APPLICATION NUMBER: US/09/595,328C  
NUMBER OF SEQ ID NOS: 3107  
SEQ ID NO 274  
LENGTH: 611  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..611  
OTHER INFORMATION: Xaa is any amino acid  
NAME/KEY: misc\_feature  
LOCATION: 1..611  
OTHER INFORMATION: Ceres Seq. ID 1009069  
US-09-595-328C-274

Query Match 21.3%; Score 863; DB 6; Length 611;  
Best Local Similarity 38.1%; Pred. No. 4.6e-61;  
Matches 203; Conservative 95; Mismatches 183; Indels 52; Gaps 15;  
QY 119 ASEQDSSEIMDANEOPQAKVTRTSIVFTGEEAAPYAKSGGLGVDGCSLPALAAARGHRVMV 178  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 123  
QY 179 MPRYLNGSSDKNAYAKALYTAHKIKIPCFGGSHEVTFEYRDNDVWVFDHPHY-----H 233  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 177  
QY 234 RPSG-LYGDNFGA-FGDNQRYTLLCYAAACEAPLILELGG--YI---YGQNCMFVNDWH 286  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 237  
QY 287 ASLPVLLAAKYRPGYVYRDSRSTLVTHNLAHQGVPEASTYPDLGLPPEWYGALEWVFE 346  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 293  
QY 347 WARRHALD-----KGEAVNFKAGVYADRTVTSOGYSWE-VTTAEGQGLNELSSR 399  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 345  
QY 400 KSVLNGIVNGIDINDNPTDKCLPHYSVDLLS-GKAKCKAEQLKELGLPVREDVPLIG 458  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 405  
QY 459 FIGRLDYQKIDILKMAIPELMREDYQVFMGLSGDPFEGWMRSTESSYKDKFRGWGFS 518  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 465  
QY 519 VPVSHRTACDILLMPSRPEPCGLNOLYAMQYGTVPVHGTGGLRDTVET-----FNP 572  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 525  
QY 573 FGAKGEGTGWASPLTVDMKMLALRTAMSTFREHKPSWEGLMKRGMTKDHWT 625  
Db SSKRSSKVKTAGKIVCEK-GMSVIFIGAEVGPWSKGTGLDGLGPPALAAARGHRVMTI 571

RESULT 14  
US-09-595-328C-276  
Sequence 276, Application US/09595328C  
GENERAL INFORMATION:  
APPLICANT: ALEXANDROV, Nikolai  
APPLICANT: BROVER, Vyacheslav  
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
FILE REFERENCE: 2750-0946P  
CURRENT APPLICATION NUMBER: US/09/595,328C  
NUMBER OF SEQ ID NOS: 3107  
SEQ ID NO 276  
LENGTH: 528  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..528  
OTHER INFORMATION: Xaa is any amino acid  
NAME/KEY: misc\_feature  
LOCATION: 1..528  
OTHER INFORMATION: Ceres Seq. ID 1009071  
US-09-595-328C-276

Query Match 21.3%; Score 862.5; DB 6; Length 528;  
Best Local Similarity 38.9%; Pred. No. 4.1e-61;  
Matches 199; Conservative 89; Mismatches 173; Indels 51; Gaps 14;











GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2002, 15:01:15 ; Search time 50.54 Seconds  
(without alignments)  
336.614 Million cell updates/sec

Title: US-09-674-824-2  
Perfect score: 4044  
Sequence: 1 MAATGVGAGCLAPSVRLRAD.....SDGSLSVRTAETIRNQLVTL 756

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2.6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2.6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2.6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2.6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2.6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3612	89.3	671	4	US-09-196-390-2
2	2749	68.0	649	4	US-09-192-909-2
3	2557.5	63.2	583	3	US-08-941-445A-13
4	2390.5	59.1	539	3	US-08-941-445A-21
5	1992	49.3	641	4	US-08-836-567-10
6	1718.5	42.5	459	4	US-08-836-567-4
7	1215	30.0	767	4	US-08-836-567-8
8	1205.5	29.8	698	3	US-08-941-445A-11
9	1184	29.3	558	4	US-08-836-567-6
10	1148.5	28.4	799	4	US-09-196-390-6
11	1058	26.4	669	3	US-08-941-445A-9
12	877	21.7	609	3	US-08-941-445A-7
13	865	21.4	533	3	US-08-941-445A-5
14	678.5	16.8	477	1	US-07-735-065-2
15	678.5	16.8	477	1	US-08-469-202-12
16	678.5	16.8	477	2	US-08-484-434C-12
17	556	13.7	677	4	US-08-836-567-2
18	556	13.7	1197	4	US-08-836-567-12
19	556	13.7	1230	2	US-08-968-542C-35
20	551	13.6	735	4	US-09-115-704-2
21	492	12.2	1674	2	US-08-968-542C-12
22	176	4.4	64	2	US-08-470-720-15
23	158	3.9	31	2	US-08-572-951-26
24	146.5	3.6	79	2	US-08-470-720-13
25	138	3.4	33	2	US-08-572-951-5
26	133	3.3	31	2	US-08-572-951-25
27	133	3.3	59	2	US-08-470-720-14

28	118.5	2.9	2205	1	US-08-093-453B-2	Sequence 2, Appli
29	114.5	2.8	666	4	US-08-961-083-2	Sequence 2, Appli
30	110.5	2.7	587	1	US-08-398-008A-2	Sequence 2, Appli
31	110.5	2.7	587	2	US-08-893-333-2	Sequence 2, Appli
32	110.5	2.7	1045	2	US-08-553-436A-6	Sequence 6, Appli
33	109.5	2.7	682	3	US-08-481-435-6	Sequence 6, Appli
34	107.5	2.7	6095	4	US-09-144-085-2	Sequence 2, Appli
35	107	2.6	1054	1	US-08-356-354-4	Sequence 4, Appli
36	107	2.6	1054	2	US-08-778-656-4	Sequence 4, Appli
37	106.5	2.6	702	4	US-09-232-200-102	Sequence 102, App
38	106.5	2.6	702	4	US-09-232-197-102	Sequence 102, App
39	106	2.6	382	3	US-08-660-645A-7	Sequence 7, Appli
40	106	2.6	382	3	US-09-298-718-7	Sequence 7, Appli
41	106	2.6	382	4	US-09-546-969-7	Sequence 7, Appli
42	106	2.6	382	4	US-08-980-832-5	Sequence 5, Appli
43	103.5	2.6	495	1	US-07-841-997A-2	Sequence 2, Appli
44	103.5	2.6	495	1	US-08-290-301-2	Sequence 2, Appli
45	103.5	2.6	495	2	US-08-588-983-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-196-390-2  
; Sequence 2, Application US/09196390  
; Patent No. 63071125  
; GENERAL INFORMATION:  
; APPLICANT: Block, Martina  
; APPLICANT: Lorz, Horst  
; APPLICANT: Lutticke, Stephanie  
; APPLICANT: Walter, Lennart  
; APPLICANT: Froberg, Claus  
; APPLICANT: Kossmann, Jens  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES  
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10020

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09196,390  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 196 21 588.9  
FILING DATE: 29-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 196 36 917.7  
FILING DATE: 11-SEP-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP97/02793  
FILING DATE: 28-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley, Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: AGREVO-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 671 amino acids



TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-196-390-2

Query Match 89.3%; Score 3612; DB 4; Length 671;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 TOPPLDAGVGLAPDLLLLBGLAEDSIDSIIVAAEODSEIMDANEQPOAKVTRISVFT 145  
DB 1 TOPPLDAGVGLAPDLLLLBGLAEDSIDSIIVAAEODSEIMDANEQPOAKVTRISVFT 60  
QY 146 GEAPYAKSGGLGDCVCSLPIALAAHRVVMVPRVYNGSSDKNAKALYTAHKKIP 205  
DB 61 GEAPYAKSGGLGDCVCSLPIALAAHRVVMVPRVYNGSSDKNAKALYTAHKKIP 120  
QY 206 FGSHEVTFPHEYRDNDVWVEVDHPSYHRPGSLYGNFGAGNQRYTLLCYAAEAPL 265  
DB 121 FGSHEVTFPHEYRDNDVWVEVDHPSYHRPGSLYGNFGAGNQRYTLLCYAAEAPL 180  
QY 266 ILEGGYIYQNCMFVNDHNSLVPVLLAAKYRPGVYRDSRSTLVHNLAHQGVPEAS 325  
DB 181 ILEGGYIYQNCMFVNDHNSLVPVLLAAKYRPGVYRDSRSTLVHNLAHQGVPEAS 240  
QY 326 TYPDLGLPPPEWYGALEWVFPPEWARRHALDKGEAVNFKGAVVTADRIVTYSQGSWEVTT 385  
DB 241 TYPDLGLPPPEWYGALEWVFPPEWARRHALDKGEAVNFKGAVVTADRIVTYSQGSWEVTT 300  
QY 386 AEGQGLNELLSRKSVLNGIVNGIDINWNPTDKCLPHYSVDDLSGKAKKAELOKE 445  
DB 301 AEGQGLNELLSRKSVLNGIVNGIDINWNPTDKCLPHYSVDDLSGKAKKAELOKE 360  
QY 446 LGIPVREDVPLIGFIRLDYQKGTDLTKMAIPELMREDVQFVMLGSDPIFEGWMRSTES 505  
DB 361 LGIPVREDVPLIGFIRLDYQKGTDLTKMAIPELMREDVQFVMLGSDPIFEGWMRSTES 420  
QY 506 SYKDKFRGWGFSVPVSHRITAGCDILLMPSPREPCGLNLYAMQYGTVPVHGTGGLRD 565  
DB 421 SYKDKFRGWGFSVPVSHRITAGCDILLMPSPREPCGLNLYAMQYGTVPVHGTGGLRD 480  
QY 566 TVETNPFAGKGBGTGWAISPLTVDKMLWALRTAMSTFHEHFKPSWGLMKRGTKDHTW 625  
DB 481 TVETNPFAGKGBGTGWAISPLTVDKMLWALRTAMSTFHEHFKPSWGLMKRGTKDHTW 540  
QY 626 DHAPSSTSRSSGSPSWNTPTSCRRGLGRSKCESPSALKTSSSSRGPEGYPCTLRCPATV 685  
DB 541 DHAPSSTSRSSGSPSWNTPTSCRRGLGRSKCESPSALKTSSSSRGPEGYPCTLRCPATV 600  
QY 686 ESQCACLLWFAGSRTYDGCAAAATAVSSGGRQLQFWGIRKCAAGWLTAHKHSDGSLSVRV 745  
DB 601 ESQCACLLWFAGSRTYDGCAAAATAVSSGGRQLQFWGIRKCAAGWLTAHKHSDGSLSVRV 660  
QY 746 TAEIRNQLVTL 756  
DB 661 TAEIRNQLVTL 671

RESULT 2  
US-09-192-909-2  
Sequence 2, Application US/09192909  
Patent No. 6307124  
GENERAL INFORMATION:  
APPLICANT: Jens Kossmann  
APPLICANT: Claus Froberg  
TITLE OF INVENTION: Nucleic acid molecules encoding soluble  
TITLE OF INVENTION: starch synthases from maize  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: James F. Haley, Jr., c/o FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York

STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/192,909  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
PCT/EP97/02527  
FILING DATE: 16-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 196 19 918.2  
FILING DATE: 17-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley, Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: GFB-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000  
TELEFAX: 212-96-9090  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 649 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-192-909-2

Query Match 68.0%; Score 2749; DB 4; Length 649;  
Best Local Similarity 82.5%; Pred. No. 1-7e-247;  
Matches 518; Conservative 34; Mismatches 62; Indels 14; Gaps 4;

QY 4 TGVGAGCLAPSVRLRADPATAARASACVWRA---RLRRLARGRYVAELSRGPAARPAQ 60  
DB 14 SAVGAACLL-----LARAAMPAAVGDRAAPRRRLQRLRRRCVAEFLSRGPAARMP 65  
QY 61 QQLAPLVPGLFAPPAPAPASPTOPPLPDAGVAGELAPDLLLEGIAEDSIDSIIVAA 120  
DB 66 ALLAPLVPGLFAPPAEPTGEPALTPPVDPDAGLVGLVE--PEGIAEGSIDNTVVVAS 122  
QY 121 EODSEIMDANEQPOAKVTRISIVFTGEAPYAKSGGLGDCVCSLPIALAAARGHRVMVMP 180  
DB 123 EODSEIVVGKEQARAKVTQNVFTGEASPYAKSGGLGDCVCSLPIALAAARGHRVMVMP 182  
QY 181 RYNGSSDKNYKALYTAHKKIPCTFGGSHSEVTFHEYRONVDVWVFDHPSYHRPGSLY 240  
DB 183 RYNGTSDKNYANAFYKHIRIPCEGGEHEVTFHEYRDSVDVWVFDHPSYHRPGSLY 242  
QY 241 DNFAGFCDNFRYTLCLCYAAACEAPLILELGGYIYQNCMFVNDHNSLVPVLLAAKYR 300  
DB 243 DKFAGPDGNFRYTLCLCYAAACEAPLVLELGGYIYQNCMFVNDHNSLVPVLLAAKYR 302  
QY 301 YGYRDSRSTLVHNLAHQGVPEASTYDGLPPEWYGALEWVFPPEWARRHALDKGEAVN 360  
DB 303 YGYVKDSRSLVHNLAHQGVPEASTYDGLPPEWYGALEWVFPPEWARRHALDKGEAVN 362  
QY 361 FLGAVVTADRIVTYSQGSWEVTTAEGGGLNELLSRKSVLNGIVNGIDINWNPTTD 420  
DB 363 FLGAVVTADRIVTYSQGSWEVTTAEGGGLNELLSRKSVLNGIVNGIDINWNPTTD 422  
QY 421 KCLPHYSVDDLSGKAKKAELOKELGPVREDVPLIGFIRLDYQKGTDLTKMAIPELM 480  
DB 423 KCLPHYSVDDLSGKAKKAELOKELGPVREDVPLIGFIRLDYQKGTDLTKMAIPELM 482  
QY 481 REDVQFVNLGSGDPIFEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSPR 540  
DB 483 REDVQFVNLGSGDPIFEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSPR 542



```
QY 541 CGLNOLYAMQYGVVPHCTGGGLRDTVETNPFGAKGEGCTGAFSPLTVDKMLALRTA 600
|||||
Db 543 CGLNOLYAMQYGVVPHATGGGURDVTENFPFENGEGCTGAFAPLTTENMLTLRTA 602
|||||

QY 601 MSTFREHKPSWEGLMKRGMTKDHTWDHA 628
|||||
Db 603 ISTRYREHKPSWEGLMKRGMSKDFTDHA 630
|||||

RESULT 3
US-08-941-445A-13
; Sequence 13, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-941-445A-13

Query Match 63.2%; Score 2557.5; DB 3; Length 583;
Best Local Similarity 86.7%; Pred. No. 1e-229;
Matches 476; Conservative 25; Mismatches 45; Indels 3; Gaps 2;

QY 45 VAEISREGPAAPAOQOOLAPPLVPGFLAPPAPPAQSPAPTOPPLPDAGVGLAPDLLLL 104
|||||
Db 2 VAEISREGPAAPRLPALLAPPLVPGFLA-PPAETGEPASTPPVPDAGLDGLG--LEP 58
|||||

QY 105 EGIADSDISIIVASEQDSEIMDANEQPAQKAVTSIVFTGEAAPIYAKSGGLGVDVCGSL 164
|||||
Db 59 EGIAGSIDNTVVASEQDSEIIVGKEQARAKVTQSIVFTGEASPIYAKSGGLGVDVCGSL 118
|||||

QY 165 PIALAARGHRVMVMPRYLNGSDSKNYAKALYKHIKIPCGGSHVTFHFHEYRDNVDM 224
|||||
Db 119 PVALAARGHRVMVMPRYLNGSDSKNYANAFYTEKHIRIPCGGGEHVTFFHFHEYRDSVDW 178
|||||

QY 225 VFVDHPSYHRPGSLYDNGFAGDNOFRYTLCLYAAACEAPLILELGGYIYGONCMFVND 284
|||||
Db 179 VFVDHPSYHRPGNLVYDNGFAGDNOFRYTLCLYAAACEAPLILELGGYIYGONCMFVND 238
|||||
```

```
QY 285 WHASLVPLVLAARKYRPGVYRDSRSLTVIHNLAHQGVPEASTYPTDLGLPPEWYGALEWVF 344
|||||
Db 239 WHASLVPLVLAARKYRPGVYKDSKSLTVIHNLAHQGVPEASTYPTDLGLPPEWYGALEWVF 298
|||||

QY 345 PEWARREHALDKGEAVNPLKGAVTADRIYTVSOGYSWEVTTAEGGQGLNELLSSRSKSVLN 404
|||||
Db 299 PEWARREHALDKGEAVNPLKGAVTADRIYTVSOGYSWEVTTAEGGQGLNELLSSRSKSVLN 358
|||||

QY 405 GIVNGIDINDNPTTDKCLPHHYSVDDLSGKAKCAELQKELGLPVRDVPPLIGFGRDL 464
|||||
Db 359 GIVNGIDINDNPTDKCIPCHYSVDDLSGKAKCAELQKELGLPVRDVPPLIGFGRDL 418
|||||

QY 465 YQKGDILIKMAPELMDREDVQFVNLGSGDPIFEGNMRSTESSYKDKFRGWGFSVPVSHR 524
|||||
Db 419 YQKGDILIQIIPDLMDREDVQFVNLGSGDPELEDNMRSTESIFKDKFRGWGFSVPVSHR 478
|||||

QY 525 ITAGCDILLMPSPFPCGLNOLYAMQYGVVPHCTGGGLRDTVETNPFGAKGEGCTGWA 584
|||||
Db 479 ITAGCDILLMPSPFPCGLNOLYAMQYGVVPHCTGGGLRDTVETNPFGAKGEGCTGWA 538
|||||

QY 585 FSPLTVDKM 593
|||||
Db 539 FAPLTTENM 547
|||||

RESULT 4
US-08-941-445A-21
; Sequence 21, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-941-445A-21

Query Match 59.1%; Score 2390.5; DB 3; Length 539;
Best Local Similarity 88.4%; Pred. No. 3.3e-214;
Matches 444; Conservative 23; Mismatches 32; Indels 3; Gaps 2;
```



```

QY 95 VGELA-PDLLL--EGTAEDSIDSLIVAASQDSEINDANEPQQAQKVRISIVFVTGSAAPY 151
Db 2 VAELSREDLGEPTAEGSIDNTVVASBQDSEIVVQKEQAKVQTQSVFVTGSAAPY 61
QY 152 AKSGGLDVCGLPIALAARGHRVWVMPRYLNGSDKNYAKALYAKHIIKICPCGSGHE 211
Db 62 AKSGGLDVCGLPVALAARGHRVWVMPRYLNGTSDKNYANAFYKHKIRICPCGSGHE 121
QY 212 VTFEFHYRDNDVWVFDVDPHPSYHRPGSLYGNFGAFGDNQFRYTLILCYAACEAPLILELGG 271
Db 122 VTFEFHYRDSVDVWVFDVDPHPSYHRPGSLYGNFGAFGDNQFRYTLILCYAACEAPLILELGG 181
QY 272 YIYGQNCMFVYNDWVHSLVPLLAARYPYGVYRDSRSTLVIHNLAHQGVPEPASTYVDLG 331
Db 182 YIYGQNCMFVYNDWVHSLVPLLAARYPYGVYRDSRSTLVIHNLAHQGVPEPASTYVDLG 241
QY 332 LPPEWYGALEWPEWARRHALDKEAVNFKGAVTADRIVTYVSGYSHEVTTAEGGQ 391
Db 242 LPPEWYGALEWPEWARRHALDKEAVNFKGAVTADRIVTYVSGYSHEVTTAEGGQ 301
QY 392 LNELLSRKSVLNGVINGIDINDWNPPTDKCLPHHYSVDDLSGKAKCKAELQKELGLPVR 451
Db 302 LNELLSRKSVLNGVINGIDINDWNPATDKCIPCHYSVDDLSGKAKCKGALQKELGLPIR 361
QY 452 EDVPLIGFICRLDYQKGDILIKMAIPELMREDVOFVMLGSDPIFEGWMRSTESSYKDKF 511
Db 362 PDVPLIGFICRLDYQKGDILQIIPDLREDVQFVMLGSDPELEDWMRSTESIFKDKF 421
QY 512 RGWVGFSVPVSHRITAGCDILLMPSPREPCGLNOLYAMQYGVVHGTGGLRDTVTETEN 571
Db 422 RGWVGFSVPVSHRITAGCDILLMPSPREPCGLNOLYAMQYGVVHGTGGLRDTVTENFN 481
QY 572 PFGAKGEGTGWAFSPILTVDKM 593
Db 482 PFGENGEGTGWAFPLTENN 503

RESULT 5
US-08-836-567-10
; Sequence 10, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens.
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836, 567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrovo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-836-567-10

Query Match 49.3%; Score 1992; DB 4; Length 641;
Best Local Similarity 67.5%; Pred. No. 6.1e-177;
Matches 360; Conservative 80; Mismatches 87; Indels 6; Gaps 4;

QY 102 LLEGTAEQDS--IDSLIVAASEQD--SEIMDANEQQAQKVRISIVFVTGSAAPYKSGGL 157
Db 90 LIPHSVAGDQATWVESHDIANDRDLSEDETEEMEETPIKLTFTNIIIFTAAPYKSGGL 149
QY 158 GDVCGSLPTALAAARGHRVWVMPRYLNGS--SDKNYAKALYAKHIIKICPCGSGSHEVTFPH 216
Db 150 GDVCGSLPVALAARGHRVWVMPRYLNGSPSDEKYNANAVDLDVRATVHCFGDAQEAFAFH 209
QY 217 EYRDNDVWVFDVDPHPSYHRPGSLYGNFGAFGDNQFRYTLILCYAACEAPLILELGGYIYGQ 276
Db 210 EYRAGVDVWVFDVDPHPSYHRPGSLYGNFGAFGDNQFRYTLILCYAACEAPLILELGGYIYGQ 269
QY 277 NCMFVNDWVHSLVPLLAARYPYGVYRDSRSTLVIHNLAHQGVPEPASTYVDLGGLPPEW 336
Db 270 KCLFLANDWHAALVPLLAARYPYGVYKARSIVAHNIHQAQGVPEPASTYVDLGGLPPEW 329
QY 337 YGALEWPEWARRHALDKEAVNFKGAVTADRIVTYVSGYSHEVTTAEGGQGLNEL 396
Db 330 YGAVWEIFPTWARAHADTGETVNLKGAITAVADRIVTYVSGYSHEVTTAEGGQGLNEL 389
QY 397 SSRKSVLNGVINGIDINDWNPPTDKCLPHHYSVDDLSGKAKCKAELQKELGLPVRDVL 456
Db 390 SSRQSVLNGITGIDVNDWNPSTDEHTASHYSINDLSGKVKQCKTDLQKELGLPIRDCPL 449
QY 457 IGFIGRDLQYKGDILIKMAIPELMREDVOFVMLGSDPIFEGWMRSTESSYKDKERGWG 516
Db 450 IGFIGRDLQYKGDVILISAPELMQNDVQVVMGSGEKQFEDWMRHTENLFKDKERAWVG 509
QY 517 FSPVSHRITAGCDILLMPSPREPCGLNOLYAMQYGVVHGTGGLRDTVTETENPFGAK 576
Db 510 FNPVSHRITAGCDILLMPSPREPCGLNOLYAMRYGTIPVHSTGGLRDTVKDFENPAAQ 569
QY 577 G-BEGTGWAFSPILTVDKMLMALRTAMSTFHEHPSWEGMLMKRGTWKDHTWDHA 628
Db 570 GIGEGTGWTFSPILTSEKLDLTALAIGTYTEHKSWSWEGMLMRGMRGDRYSWENA 622

RESULT 6
US-08-836-567-4
; Sequence 4, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York

```







Db 478 GEHNFIAAGLTADRVVTVSHGYSWELKTSQGGMLHGOIINENDKWLGQIVNGIDTKEW 537  
Qy 416 NPITDKCLPH---HYSVDL-SGKAKCAELQKELGLPVRDVPVPLIGFIRLDYQKGD 470  
Db 538 NPELDVHLQSDGVMNLSLDTQTKPKQCAALQKELGLPVRDVPVPLIGFIRLDYQKGD 597  
Qy 471 LIKMAPELMREDYQVFMVLMGSDPIFEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCD 530  
Db 598 LIAEASAMWQDVLVMLTGRDLEQMLRQFECQHNDKIRGWGFSVKTSHRITAGAD 657  
Qy 531 ILLMPSFEPCGLNOLYAMQYCTVPVHGTGGLRDVTETNPEFGAKGEGTGWAFSPLTV 590  
Db 658 ILLMPSFEPCGLNOLYAMQYCTVPVHGTGGLRDVTETNPEFGAKGEGTGWAFSPLTV 590  
Qy 591 DKMLWALTAMSTFREHKPWSWGLMKRGMTKDHWDHA 628  
Db 715 SOLIHALGNCLTYREYKKSMEGIQTRCMTQDLSWDA 752  
RESULT 8  
US-08-941-445A-11  
: Sequence 11, Application US/08941445A  
: Patent No. 6107060  
: GENERAL INFORMATION:  
: APPLICANT: Keeling, Peter  
: TITLE OF INVENTION: Starch Encapsulation  
: NUMBER OF SEQUENCES: 37  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
: STREET: 5370 Manhattan Circle  
: CITY: Boulder  
: STATE: CO  
: COUNTRY: US  
: ZIP: 80303  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/941.445A  
: FILING DATE: 30-SEP-1997  
: CLASSIFICATION: 800  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 60/026,855  
: FILING DATE: 30-SEP-1996  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Winner, Ellen P  
: REGISTRATION NUMBER: 28,547  
: REFERENCE/DOCKET NUMBER: 89-97  
: TELEPHONE: (303) 499-8080  
: TELEFAX: (303) 499-8089  
: INFORMATION FOR SEQ ID NO: 11:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 698 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US-08-941-445A-11  
Query Match 29.8%; Score 1205.5; DB 3; Length 698;  
Best Local Similarity 43.0%; Pred. No. 1.4e-103;  
Matches 283; Conservative 83; Mismatches 221; Indels 71; Gaps 17;  
Qy 12 APSVRLRADPA-----TAASACVVRRLRLRGRVYAELSREGPARPAQOQLAPP 66  
Db 56 AASVRAAAAPAGSEAEAKSSS-SQAGAVQGSTAKAVDSASPPNPLTSPAKQSQRAM 114  
Qy 67 L--VPGFLAPPPAPPAQSPAPTOPPLP-----DAG-----VGLAPDLLLEGIAEDSI 112

115 QNGTSGSSASTAAPVSGPKADHPSAPVTKREIDASAVKPEAGDDARPVESIGIAE-PV 173  
Qy 113 DSIIVAASEQDS-----EIDANEQQA-KVTRSIYFVTGEAAPYAKSGGLGDDVCGSLP 165  
Db 174 DAKADAAPATDAASAPYDREDNEPGLAGNVMNVMNVVVAASECAPCKTGGLGDDVVGALP 233  
Qy 166 IALAARGHRVWVMPRYLNGSSDKYAKALYAKHIKIPCFGSGSHSEVYTFTHYRDNDVWY 225  
Db 234 KALARGHRVWVMPRY-----GEYAEARDLGVRRYKAVAGQDSVYTFHSHYIDGVDFV 287  
Qy 226 FVDHPSY-HRPGSLYGDNFGAGNQFRYTLCLCYAAEAPLILELGGYIYCO-NCMEVNV 283  
Db 288 FVEADPERHRRNNIYG---GERLDILKRMILFEKAAVEVPWYAPCGGTGVYGDGDLVFIA 344  
Qy 284 DWHASLVVLLAAKYRYPYGVYRDSRSTLVIHNLHQGVPEASTYPDLCGLPPEWYGALEW 343  
Db 345 DMHTALLPVYKAYYRDNGLMQYARSVLVHIAHQHGRGPVDDFVNFDEL----- 394  
Qy 344 FPEWARRHALDK-----GEAVNFKGAVVTVADRVTVVTSQGYSWETTTAGGQGLNEL 395  
Db 395 -----EHYIDHFKLYDNIGGDHNSVFAAGLKTADRVTVVTSVNGYMWELKTSEGGMLHDI 448  
Qy 396 LSSRKSLVINGIDINDWNPPTDKCLPH---HYSVDL-SGKAKCAELQKELGLP 450  
Db 449-INONDKLQGIYVNDISEWNPVAVDVHLSHDDYTYTFETLDTGKQCKAALQRLGLQV 508  
Qy 451 REDVPLIGFIRLDYQKGDILKMAIPELMREDVQFVLMGSDPIFEGWMRSTESSYKDK 510  
Db 509 RDDVPLIGFIRLDYQKGDILKMAIPELMREDVQFVLMGSDPIFEGWMRSTESSYKDK 568  
Qy 511 FRGWGFSVPVSHRITAGCDIILLMPSRPEPCGLNOLYAMQYCTVPVHGTGGLRDVTET 570  
Db 569 VRAWGFSVPLAHRITAGADILLMPSRPEPCGLNOLYAMQYCTVPVHGTGGLRDVTET 628  
Qy 571 NPEGAKGEGTGWAFSPLTVDKMLWALTAMSTFREHKPWSWGLMKRGMTKDHWDHA 628  
Db 629 DPF---NDTGLGWTDFDRAENRMIDALSCLTTYRNYKESWRACRARGMAEDLSMDHA 683  
RESULT 9  
US-08-836-567-6  
: Sequence 6, Application US/08836567  
: Patent No. 6130367  
: GENERAL INFORMATION:  
: APPLICANT: Kossmann, Jens  
: APPLICANT: Springer, Franziska  
: APPLICANT: Abel, Gernot  
: TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES  
: TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC  
: TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: FISH & NEAVE  
: STREET: 1251 Avenue of the Americas  
: CITY: New York  
: STATE: New York  
: COUNTRY: USA  
: ZIP: 10020  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/836,567  
: FILING DATE: 24-JUL-1997  
: CLASSIFICATION: 800  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: PCT/EP95/04415  
: FILING DATE: 09-NOV-1995  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: DE P 44 41 408.0  
: FILING DATE: 10-NOV-1994







Db	386	DFVFDAPLFRHQRQDIYG---GSRQEIIMKRMILFCKAAVEVPHWVPCGGVYGDGNLVE	442
Qy	281	VVNDWASLVVPIVLAARYPYGYRDSRSTLVVHNLAHQGVPEASTYVPLGLPPEWYGAL	340
Db	443	IANDWHTALLPVYLKAYYRDHGLMQVTRSMVTHNIAHQGRGVPDEFPELPEHYL---	499
Qy	341	EWVPEWARHALDKGEAVNFKGAVVADRIYTVVSQGYSWEVTTAEGGGLNELLSSRK	400
Db	500	-----EHFRLYDPVGEHANYFAAGLKMDQVVVVSPGYLWELKTVEGGWGLHDIIRQND	554
Qy	401	SVLNGIVNGIDINDWNPPTTKCLPH-----HYSVDDL-SGKAKCKABLOKELGLPVREDVP	455
Db	555	WKTGIVNGIDNMENPEVDHAKLSQDGTNFSRLTLDGSKRQCKEALQRELGLQVRADVP	614
Qy	456	LIGFIGRLDYOKGIDLIKMAIPELMRDQVFMVLMGSDPIFEGWMSRSTESSYKDKRGGW	515
Db	615	LLGFIGRLDGQKGVEITADAMPWIVSQDVQLVMLGTGRHDLSEMLQHFEREHDKVRGW	674
Qy	516	GFSVPVSHRITAGCDILMLPSRPEPCGLNOLYAMQYGVTPVPHGTGRLDRTVETFPFGA	575
Db	675	GFSVRLAHRITAGADALLMSRPEPCGLNOLYAMAYGTVPVHVAVGGLRDTVPPEDPF---	732
Qy	576	KGEEGTGFWSPLTVDKMLMALRTAMSTFREHKPSWEGMLMKRMTKDHTWDHA	628
Db	733	NHSLGCLWTFDRAEAHKLIEALGHLCLRTYRDFKESWRALQERGMSQDFSWEHA	784
RESULT 11			
US-08-941-445A-9			
; Sequence 9, Application US/08941445A			
; Patent No. 6107060			
; GENERAL INFORMATION:			
; APPLICANT: Keeling, Peter			
; APPLICANT: Guan, Hanping			
; TITLE OF INVENTION: Starch Encapsulation			
; NUMBER OF SEQUENCES: 37			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.			
; STREET: 5370 Manhattan Circle			
; CITY: Boulder			
; STATE: CO			
; COUNTRY: US			
; ZIP: 80303			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patent In Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/941.445A			
; FILING DATE: 30-SEP-1997			
; CLASSIFICATION: 800			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 60/026,855			
; FILING DATE: 30-SEP-1996			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Winner, Ellen P			
; REGISTRATION NUMBER: 28,547			
; TELEPHONE: (303) 499-8080			
; TELEFAX: (303) 499-8089			
; INFORMATION FOR SEQ ID NO: 9:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 669 amino acids			
; TYPE: amino acid			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
US-08-941-445A-9			
Query Match 26.4%; Score 1068; DB 3; Length 669;			

Best Local Similarity 38.0%; Pred. No. 8.6e-91;			
Matches 256; Conservative 99; Mismatches 236; Indels 82; Gaps 17;			
Qy	12	APSVRLRADPATAARASACVVRARLRRLAR--GRY---VAELSREGPAARPAQOOLAP--	65
Db	10	APPER--SGDAARLPARRNNAVSKRRDPLQPVGRYGSATGNTARTGAAS--CQNAALADVE	66
Qy	66	-PLVPGFLAPPPAPAPQSPAPTQPLPDAGVGLAPDLLL-----EGIA	108
Db	67	IVEIKSIVAAPTSTVTKRPG--RGLQDDPSLWDIAPETVLPAKPLHESPAVDGDSNGIA	124
Qy	109	EDSIDSIIVAAEQSQSEIMDANEQPAK-----VTRSIIV	142
Db	125	PPTVEPLVQEAATWDFKTYTGFDEPDEAKDDSRVAGDADGASPEHYMTMILGCGENVMVI	184
Qy	143	FVTGGAAPYAKSGGLDVCGLPIALAAARHVRVMVMPYRNGSSDKNYAKALYTAHIK	202
Db	185	VVAACECPCKTGGGLGDVVGALPKALARHVRVMVVPYR-----GDYVEAFDMGIRKY	238
Qy	203	IPCFGGSHEVTFHEYRDWVWVVDHPSYHRPGSLYGDNFCAGDGNQFRYTLLCYACE	262
Db	239	YKAAGODLEVNYPFAFIDGVDFEVIDASPRHRQDDIYG---GSRQEIIMKRMILFCKVAVE	295
Qy	263	APLILELGGYIYGO--NCMEFVNDWHASLVPLLAARYPYGYRDSRSTLVVHNLAHQGV	321
Db	296	VMHVPCCGVCYGDGNLVFIAMNHTALLPVYLKAYYRDHGLMQYTRSLVLIHNIGHOGR	355
Qy	322	EPASTYPDILGLPPEWYGALEWVFPPEWARRALDKGEAVNFKGAVV--TADRIYTVVSQGY	380
Db	356	GPVHEFPYDMLNTNLQHFELYDPV-----GGSHANIFAACVLKMDRVTVTSRGLY	407
Qy	381	WEVTTAEGGGLNELLSSSRKSVLNGIVNGIDINDWNPPTTKCLPH-----HYSVDDL-SCK	435
Db	408	WELKTVEGGWGLHDIIRSDNWKINGIRERIDHQEWPKYDVHLRSDGYTNYSLTLDAGK	467
Qy	436	AKCAEALQELGLPVREDVPLIGFIGRLDYOKGIDLIKMAIPELMRDQVFMVLMGSDPI	495
Db	468	RQCAALQRDVGLVLEVRDVPVPLIGFIGRLDGQKGVDIIGDAMPWIAQODVQLVMLGTGPPD	527
Qy	496	FEQWKRSTESSYKDKFRGWVGFSPVSHRITAGCDILLMPSPRPEPCGLNOLYAMQYGTVP	555
Db	528	LERMLQHLEREHPNKVRGWVGFSLVMVHRITPGASVLVMPSPFAG--GLNOLYAMAYGTVP	586
Qy	556	VHGTGGLRDTVETFPFGAKGEGTGWAFSLTVVDKMLMALRTAMSTFREHKPSWEGML	615
Db	587	VHAVGGLRDTVAFDPFF---GDAGLWTFDRAEANKLIEVLSHCLDTYRNYEESNKSQ	643
Qy	616	KRGMTKDHTWDHA	628
Db	644	ARGMSQLNSWDHA	656
RESULT 12			
US-08-941-445A-7			
; Sequence 7, Application US/08941445A			
; Patent No. 6107060			
; GENERAL INFORMATION:			
; APPLICANT: Keeling, Peter			
; APPLICANT: Guan, Hanping			
; TITLE OF INVENTION: Starch Encapsulation			
; NUMBER OF SEQUENCES: 37			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.			
; STREET: 5370 Manhattan Circle			
; CITY: Boulder			
; STATE: CO			
; COUNTRY: US			
; ZIP: 80303			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patent In Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/941.445A			
; FILING DATE: 30-SEP-1997			
; CLASSIFICATION: 800			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 60/026,855			
; FILING DATE: 30-SEP-1996			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Winner, Ellen P			
; REGISTRATION NUMBER: 28,547			
; TELEPHONE: (303) 499-8080			
; TELEFAX: (303) 499-8089			
; INFORMATION FOR SEQ ID NO: 9:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 669 amino acids			
; TYPE: amino acid			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
US-08-941-445A-9			
Query Match 26.4%; Score 1068; DB 3; Length 669;			



```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-941-445A-7

```

```

Query Match 21.7%; Score 877; DB 3; Length 609;
Best Local Similarity 38.5%; Pred. No. 4.7e-73;
Matches 205; Conservative 89; Mismatches 173; Indels 66; Gaps 15;

QY 140 SIIVFTGEAAPAKSGGLGDCVGLSLPILAAARGHRVVMVPRYLNGSSDKNYAKALYTAK 199
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 84 NVVFGAEMAPWSKTGLGDLVGLPPAMAANGHRVVMVSPRY-----DQYKADWDTVS 137
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 200 HKIPCFGSGHEVTFPHFYRDNDVWFVDHPSY-----HRPG-SLYGDNFGA-FGDNQFR 252
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 138 VAEIKVADRYERVRFFHCYKRGVDRFIDHPSFLEKRVWCKTEGKIYGPDTGVYKDNQMR 197
QY 253 YTLICVAACEAPLILELGGY-----IYQNCMFVNDWHASLVPLVLLAAKYRYPYVRS 307
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 198 FSLCCQAALAPRILNLLNNPFYKGTGYEDVVFVNCNDWHTGTLASLYLKNYQPNGIYRNA 257
QY 308 RSTLVIIHLAHOGVEPASTYDPLGLPPEWYGALEWVPFWARRHALD---KGEAVNFKG 364
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 258 KVAFCIHNIYQGRFAFEDYPELNLSEFRSSDFI-----DGYDTPTVEGRKINWKA 310
QY 365 AVVTADRVTVSQGSYWE-VTTAEGGGQGLNELLSRKSRLVINGIDINDNPTDCKL 423
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 311 GLEADRVLTSPYAEELISGIARGCELDNIM--RLTGITGIVNGMDVSEWDPKDYI 368
QY 424 PHHY-SVDDLSGKAKCAELQELGLPVREDVPLIGFIGRLDYOKGIDLIKKAIPELMRE 482
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 369 TAKYDATTAEAKALNKEALQELGLPVREDVPLIGFIGRLDYOKGIDLIKKAIPELMRE 428
QY 483 DVQFVNLGSGDPIFEGWMRSTESSYKDKFRGWGVFSPVSHRITAGCDILLMPSPFCG 542
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 429 DVQVILLGTGKKKFEKLLKMEKYPGKRVAVVKFNAPLAHLIMAGADVLAVPSRPECG 488
QY 543 LQOLYAMQYTPVVGHTGGLRDTVETFPNFGAKGEEGTGAFSPLTVD-----K 592
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 489 LIQOLQMRGTGTPCACASTGGLVDVTI-----EGKTGFHMGRLSDCKVYVPSDVAK 539
QY 593 MLWALRTAMSTFREHKPSWEGMLKMGTKDHTWDHAPSTSRSSGP---SWTN 643
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 540 VAATLKRAIKVV-GTPAYEEVVRNMCNQDLSW-----KGPKNWEN 579

```

```

RESULT 13
US-08-941-445A-5
; Sequence 5, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 533 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-941-445A-5

Query Match 21.4%; Score 865; DB 3; Length 533;
Best Local Similarity 38.0%; Pred. No. 5e-72;
Matches 205; Conservative 90; Mismatches 177; Indels 68; Gaps 16;

QY 135 AKVTSIVTGEAAPAKSGGLGDCVGLSLPILAAARGHRVVMVPRYLNGSSDKNYAKA 194
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 195 LYTAHKIKIPCFGSGHEVTFPHFYRDNDVWFVDHPSY-----HRPGSLYGDNFGA-FG 247
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 55 WTSVSVSEILKMGDGYETVRFHCYKRGVDRFVDHPLFLERVWGTETKIYGPVAGTDYR 114
QY 248 DNOFRTYLLCYAACERAPLILELGG--YI---YQNCMFVNDWHASLVPLVLLAAKYRYPY 302
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 115 DNOLRFLSLCCQAALAPRILNLLNNPFYKGTGYEDVVFVNCNDWHTGTLASLYLKNYQ 174
QY 303 VYDRSRSTLVIIHLAHOGVEPASTYDPLGLPPEWYGALEWVPFWARRHALD---GEAV 359
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 175 IYRDAKTAFCIHNIYQGRFAFSDYPELNLSEFRSSDFI-----DGYEKPVEGRKI 227
QY 360 NFKGAVVTADRVTVSQGSYWE-VTTAEGGGQGLNELLSRKSRLVINGIDINDNPT 418
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 228 NNMKAGILEADRVLTSPYAEELISGIARGCELDNIM--RLTGITGIVNGMDVSEWDP 285
QY 419 TDKCLPHHYSVDD-LSGKAKCAELQELGLPVREDVPLIGFIGRLDYOKGIDLIKKAIP 477
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 286 ROKYIAVKYDVSTAVEAKALNKEALQAEVGLPVDNRNIPLAVFIFGLEEQGPDVMAAIAIP 345
QY 478 ELMR--EDVQFVNLGSGDPIFEGWMRSTESSYKDKFRGWGVFSPVSHRITAGCDILLMP 535
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 346 QLMEMVEDQVILLGTGKKKFEKLLKMEKYPGKRVAVVKFNAPLAHLIMAGADVLAVT 405
QY 536 SRFEPCGLNQLYAMQYTPVVGHTGGLRDTVETFPNFGAKGEEGTGAFSPLTVD----- 591
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 406 SRFEPCGLLIQOMRYGTGTPCACASTGGLVDVTI-----EGKTGFHMGRLSDCKV 456
QY 592 -----KMLWALRTAMSTFREHKPSWEGMLKMGTKDHTWDHAPSTSRSSGP---SWTN 643
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```



Db 457 EPADVKKVATTLOAIAKV--GTPAYEEMVRNMCMIQDLSW-----KCPAKNWN 503

RESULT 14

US-07-735-065-2

Sequence 2, Application US/07735065

Patent No. 5349123

GENERAL INFORMATION:

APPLICANT: Christine K. Shewmaker

APPLICANT: David M. Stalker

TITLE OF INVENTION: Glycogen Biosynthetic Enzymes

TITLE OF INVENTION: In Plants

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Calgene, Inc.

STREET: 1920 Fifth Street

CITY: Davis

STATE: CA

COUNTRY: USA

ZIP: 95616

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Macintosh 6.0.7

SOFTWARE: Microsoft Word 4.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/735,065

FILING DATE: 19910724

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/632,383

FILING DATE: 21-DEC-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: unassigned

FILING DATE: 16-JUL-1991

ATTORNEY/AGENT INFORMATION:

NAME: Elizabeth Lassen

REGISTRATION NUMBER: 31,845

NAME: Donna E. Scherer

REGISTRATION NUMBER: 34,719

REFERENCE/DOCKET NUMBER: CGNE 77-2

TELEPHONE: (916) 753-6313

TELEFAX: (916) 753-1510

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 477 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-07-735-065-2

Query Match 16.8%; Score 678.5; DB 1; Length 477;

Best Local Similarity 34.1%; Pred. No. 1e-54;

Matches 170; Conservative 78; Mismatches 210; Indels 41; Gaps 12;

QY 141 IVFVTGEAPYAKSGGLGDCVCSLPTALAAARGHVRVWVMPRY---LNGSSDKNYAKALYT 197

Db 3 VLVHCSEMFLLTKGLADYIGALPAAQIADGVDAVLLPAFPDIRGVTDQVVSRRDT 62

QY 198 AKHIKIPCGSGSEVTFHFHYRONVDWVFDHPS-YHRPGSLYGD-NFGAFGDNQFRYTL 255

Db 63 -----FAG--HITLLFGHYNGVGYLLDAPHLYDRPGSPYHDTNLFAYTDNVLRAL 112

QY 256 LCVAACEAPLILEGGVIYQNCVFVNDWNASLVVLLAAKYRPGYVYRDSRSTLIHN 315

Db 113 LGHWGAE-----MASGLDPFRPVDVHHADHAGLAPAYLAARGP-----AKSVFTGHN 162

QY 316 LAHOGVEPASTYFDLGLPPEWYGALEKVFPEARRHALDKGEAVNFKLGAVVTADRTVTV 375

Db 163 LAYQGMFYAHMNDIQLP-----WSF-----FNIHGLEFNGQISFLKAGLYADHITAV 211

QY 376 SQGISHEVTTAEGCGQLNELLSS--RKSVLNGIVNGIDINDWNPPTDKCLPHHSYVDDLS 433

Db 212 SPTYAREITEPQFAYGMEGLLQQRHREGRLSGVLNGVDEKINSIPETDILLASRYTRDTLE 271

QY 434 GKAKCKAEQLKELGFPVREDVPLIGFTIGRLDYQKGLDLIKMAIPELMREDVQFVMLGSGD 493

Db 272 DKAENKRLQIANGLVDDKVPFAVVRSLTSQGLDLVLEALPGLLEGGQLALLGAGD 331

QY 494 PIFEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMPSRFPFCGLNQLYAMQYGT 553

Db 332 PVLQEGFLAAAEYPGQVGVOIGYHEAFSHRIMGADVILVPSRFPFCGLTQLYGLKYGT 391

QY 554 VPVVHGTGGLRDIVETENPEFGARGEGBTGWAFSPFLVDKMLMALRTAMSTFRSHKPS-WE 612

Db 392 LPLVRRRTGLADTVSDCSLENLADGVASGEVFDENSNWLLRAIRAFVLM--SRPSLMR 449

QY 613 GLMKRGMTKDHTWDHAPSS 631

Db 450 FVORQAMDFSWQVAKS 468

RESULT 15

US-08-469-202-12

Sequence 12, Application US/08469202

Patent No. 5750875

GENERAL INFORMATION:

APPLICANT: STALKER, DAVID

TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: CALGENE, INC.

STREET: 1920 FIFTH STREET

CITY: DAVIS

STATE: CA

COUNTRY: USA

ZIP: 95616

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Macintosh 7.1

SOFTWARE: Microsoft Word 5.1 (a)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,202

FILING DATE: 6-JUNE-95

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/016,881

FILING DATE: 11\_FEB\_1993

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Elizabeth Lassen

REGISTRATION NUMBER: 31,845

NAME: Donna E. Scherer

REGISTRATION NUMBER: 34,719

NAME: Carl J. Schwedler

REGISTRATION NUMBER: 36,924

REFERENCE/DOCKET NUMBER: CGNE 93-2

TELEPHONE: 916-753-6313

TELEFAX: 916-753-1510

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 477 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-08-469-202-12

Query Match 16.8%; Score 678.5; DB 1; Length 477;

Best Local Similarity 34.1%; Pred. No. 1e-54;

Matches 170; Conservative 78; Mismatches 210; Indels 41; Gaps 12;



Search completed: March 28, 2002, 16:21:44  
Job time: 4829 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: March 28, 2002, 12:08:26 ; Search time 2297.02 Seconds  
(without alignments)  
12963.123 Million cell updates/sec

Title: US-09-674-824-1  
Perfect score: 2771  
Sequence: 1 cgcacactccactgcctctgc.....aaaaaaaaaaaaaaaaaaaaa 2771

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hic:\*  
10: gb\_est1:\*  
11: gb\_est2:\*  
12: gb\_hic:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rod:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	601	21.7	988	11	BG343782 HVSMG000
C 2	564	20.4	706	10	AV832913 HVSMG000
3	494	17.8	988	10	BE418354 SCL031.D0
C 4	403.8	14.6	557	11	BG607242 WHE2493_B
5	393	14.2	562	10	AM678087 WSL_13.G1
6	388	14.0	454	10	AV832689 AV832689
7	380.8	13.7	408	11	BE482660 WHE2301-2
8	325.6	11.8	596	11	BE113316 BE113316
9	320.2	11.6	588	10	BE434991 BE434991
10	294.2	10.6	396	10	AU075475 AU075475
11	276	10.0	401	11	BG050773 FM1_70.D0
12	250.6	9.0	564	10	AW932199 EST358042

13	225.2	8.1	618	10	AW759569
14	223.4	8.1	623	10	AW759946
15	216.4	7.8	335	11	C19776
16	215.2	7.8	411	10	AW438020
17	205	7.4	217	11	BF478674
18	195.4	7.1	592	10	AW593325
19	194.8	7.0	294	10	AW418330
20	182.2	6.6	243	10	AA749940
21	168.6	6.1	548	11	BF424294
22	165.2	6.0	601	11	BG453184
23	159	5.7	632	10	AW678169
24	156	5.6	729	11	BG599615
25	154.8	5.6	773	11	BG599615
26	154.4	5.6	696	11	BG351175
27	148.8	5.4	734	11	BG351920
28	145.8	5.3	625	10	BE347900
29	145.4	5.2	649	11	BI406798
30	144.2	5.2	455	10	BE555965
C 31	141.4	5.1	143	10	BE425099
32	141.2	5.1	753	11	BF460244
33	140	5.1	589	10	AW569428
C 34	136.8	4.9	614	10	AV524142
35	136.2	4.9	446	11	BG647513
36	131	4.7	568	11	BG350144
37	130	4.7	679	11	BI406287
38	128.6	4.6	522	11	BG652154
C 39	127.4	4.6	754	11	BG351232
40	126	4.5	565	11	BG154708
C 41	125.2	4.5	635	10	AV441053
42	121.6	4.4	555	10	AW472190
43	121.2	4.4	794	11	BI406951
44	120.8	4.4	681	11	BG456652
45	120.4	4.3	665	11	BI406876

#### ALIGNMENTS

RESULT 1

BG343782

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BG343782 988 bp mRNA EST 29-MAY-2001  
HVSMEG0006L20f Hordeum vulgare pre-anthesis spike EST library  
HVCN00008 (white to yellow anther) Hordeum vulgare cDNA clone  
HVSMEG0006L20f, mRNA sequence.  
BG343782  
BG343782.1 GI:13156111  
EST.  
barley.  
Hordeum vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
; Triticeae; Hordeum.  
1 (bases 1 to 988)  
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu  
, Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo  
, T., Saski, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Main, D. and  
Wood, T.  
Development of a genetically and physically anchored EST resource  
for barley genomics  
Unpublished (2000)  
Contact: Wing RA  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: AATTAACCTCACTAAAGG  
High quality sequence stop: 845.  
Location/Qualifiers  
1..988  
/organism="Hordeum vulgare"  
/cultivar="Morex"



LOCUS AV832913 706 bp mRNA EST 22-JUN-2001  
DEFINITION AV832913 K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare top three leaves adult, heading stage Hordeum vulgare subsp. vulgare cDNA clone rbaal2b23, mRNA sequence.  
ACCESSION AV832913  
VERSION 1 GI:14525002  
KEYWORDS EST.  
SOURCE Hordeum vulgare subsp. vulgare.  
ORGANISM Hordeum vulgare subsp. vulgare.  
REFERENCE Sato, K.  
AUTHORS Barley EST sequencing project in NIG and Okayama Univ  
TITLE Unpublished (2001)  
JOURNAL Contact: Kazuhiro Sato  
COMMENT Research Institute for Bioresources  
Okayama University, Barley Germplasm Center  
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan  
Email: kazsato@rib.okayama-u.ac.jp/  
URL: http://www.rib.okayama-u.ac.jp/barley/  
Sato, K., Saisho, D., Takeda, K., Shini, T. and Kohara, Y. Direct  
submission:  
database: http://www.shigen.nig.ac.jp/barley/barley.html.

BASE COUNT 243 a 223 c 291 g 211 t 20 others  
ORIGIN

Query Match 21.7%; Score 601; DB 11; Length 988;  
Best Local Similarity 91.1%; Pred. No. 2e-52;  
Matches 704; Conservative 0; Mismatches 59; Indels 10; Gaps 6;  
Qy 1787 ccgagtcgagttacaagataaaattccgtggtggtggtgatttagttccagtttccc 1846  
Db 1 CCGAGTCGAGCTACAAGGATAAATTCCTGGATGGTGGATTAGTGTTCAGTTTCCC 60  
Qy 1847 acagaataactcgaggtgagatattgtaatgccatcgagatttgaaccttgcggtc 1906  
Db 61 ACAGAATAAATCGAGGTGCGATATATGTTAATGCCATCCAGATTGAACCTTCCGGTC 120  
Qy 1907 ttaacagctatgctatgcaataggtacagttctctgtagttcatggaactgggggcc 1966  
Db 121 TTAATCAGCTATATGCTATGCAATATGGTACAGTTCCTGTTGTTTCATGGAACGTGGGGCC 180  
Qy 1967 tccgagacacagtcgagaccttcaaccttttgggtgcaaaaggagagaggtacaggg 2026  
Db 181 TTCGAGACACGGTGAGACCTTCAACCTTTTGGTGTGCAAAAGGAGAGGGGTACAGGGT 240  
Qy 2027 gggggttcaccgtaccgtgacagaagtgttggtggcattgcgaacgcgagtcgca 2086  
Db 241 GGGGGTTCACACCTTCAACCTGGTGAATAAATGTTGTGGCATTCGGAACCCGATTTCCA 300  
Qy 2087 cattcaggagacacagcgtccctggggggggtcctgaagcggaggtacgacgaagacc 2146  
Db 301 CGTTTAGGGAGACACAGCCCTCCTGGGAGGGGCTCATGAAGCGAGGATGACGAAGACC 360  
Qy 2147 atactgggaccatgccc-cgagacagtcagagagatcttcgagtggtggccttcgtgacc 2205  
Db 361 ATACGTGGGACCATGCCGCCGAGCAGTACGAGCAGATCTTCGAGTGGGCCCTTCGTGGACC 420  
Qy 2206 aacctctacgtatgacgagggactgggaggtccaaagtcgagtcgagtccttcagctctg 2265  
Db 421 AACCTACGTATGATGAGACTGGGACTGGGGAGGTGCAAGCCGCCAGTCCTTGAGCTCTG 480  
Qy 2266 aagacatcctcttccttcctcgccggcccgga---aggataccctgtacattcggtgt 2322  
Db 481 AAGACATGTTCTCTCATCTTCGCGGCCCGGAAGGAGGATACCCCTGTACATTGCGTTGT 540  
Qy 2323 cctgtacagtagagtcgaatgagccttctgttctgttgggtcgccggttcgagaacatat 2382  
Db 541 CCTGCTACAGTAGAGTGCATTCGCTTGTGTC--TTGGTTTCGGCGGTTCGAGAGTAGAC 599  
Qy 2383 qacgactgtgctgctcgccggtgacagcttcgggtgacacag-tcacagttttggg 2441  
Db 600 GACGTCGCAACCTTCGCCCGG---CAGCTTCGGGTGGATGACAGNNTACAGTTNTGGGG 656  
Qy 2442 aataaggnaaggatgtgctgcaggatggttaacagcaagcaccactcagatggcagc-c 2500  
Db 657 AATAAGGAAGGGAGTGTGTCAGGATGNTGACAGAAAGACACACCTCAGTCGCAAGCCTC 716  
Qy 2501 tctctgctcggtgttacagctgaataacgaacaaactgtgactcttagcct 2553  
Db 717 TCTCGGTGGGGGTACAGCTGAATCAGAAACCCACTGGTGTCTNCTTAGCCT 769

RESULT 2  
AV832913/c

FEATURES  
source

BASE COUNT 174 a 213 c 162 g 155 t 2 others  
ORIGIN

Query Match 20.4%; Score 564; DB 10; Length 706;  
Best Local Similarity 91.9%; Pred. No. 1.2e-48;  
Matches 651; Conservative 0; Mismatches 47; Indels 10; Gaps 5;

Qy 1914 gctatatgctatgcaataggtacagttcctgtagttcatggaactggggccctccgaga 1973  
Db 706 GCTATATGCTATGCAATATGCTACAGTTCCTGTTGTTTCATGGAACCTGGGGCCTTCGAGA 647  
Qy 1974 cacagtcgagaccttcaaccttttgggtgcaaaaggagagaggtacaggggtggcgctt 2033  
Db 646 CACGCTGGAGACCTTCAACCTTTTGGTGTGCAAAAGAGAGAGGGGTACAGGGTGGCGGTT 587  
Qy 2034 ctcaccgctaacctgggacagatgttgggttcgcaacccgagtcgacacattcag 2093  
Db 586 CTCACCACTAACCTGCAAAAAATGTTGTGGCATTCGGAACCGCATTCGACGTTTAG 527  
Qy 2094 ggaacacagccgtccctgggaggggctcagaagcagaggtacgaacacacacacacac 2153  
Db 526 GGAGCAACAGCCCTCCCTGGGAGGGGCTCATGAGCGAGGATGACGAAGACCATACATGCTG 467  
Qy 2154 ggacatgtcc-ccgagcagtcacgagcagacacacacacacacacacacacacacac 2212  
Db 466 GCACCATGCCGCCGAGCAGTACGAGCAGATCTTCGAGTGGGCCCTTCGTGGACCAACCTCA 407  
Qy 2213 cgtcatgtagcaggggactggggagggtccaaagtcgagtcgagtccttcagctcctgaagcat 2272  
Db 406 CGTCATGTAGACTGGGACTGGGGAGGTCGAAGCGCCAGTCTCCCTTGAGCTCTGAAGACAT 347  
Qy 2273 cctcttcactcctccgagcccgga---aggataccctgtacattcggttgcctcacta 2329  
Db 346 GTTCTCTCATCTTCGCGGCCCGGAAGGAGGATACCCCTGTACATTGCGTTGCTCTGCTA 287  
Qy 2330 cagtagagtcgcaatgcgcctgctgtgttgggttcgcgggttcgagacacacacacacacac 2389



```

|||||
Db 286 CAGTAGAGTCGAATGCGCTTGCTGC-TTGGTTCCGCGGTTCGAGAGTAGACACGCGTC 228
|||||
QY 2390 gtgctgctgcgcgcgtgacagcttcggtgagacagcttttacagcttttggggaataagga 2449
|||||
Db 227 GCAACGCTGCCCGCGG---CAGCTTCGGGTGGATGACAGTTACAGTTTGGGGATAAGGA 171
|||||
QY 2450 agggatgtctgcaggtggttaacagcaaaagcaccactcagatggcagctctct--gt 2507
|||||
Db 170 AGGGAGGTGTCGAGGATGTTGACAGCAAGACCACTCAGTCGGCAGGCTCTCTCGGT 111
|||||
QY 2508 ccgtgttacagctgaaatcagaacacaaactggtgactcttttagccttagtgattgtgaag 2567
|||||
Db 110 GGGTGTTACAGCTGAATCAGAAACCAACTGGTCTCTTTAGCTTAGTGATGATGAAG 51
|||||
QY 2568 ttgtgacctctgtatgttctgtctcttagctgacaaatattt 2615
|||||
Db 50 TTTGTTGCCCTCTGTGTATGTTATCTTGTCTTGTAGCTNATNATAATTT 3

RESULT 3
BE418354 988 bp mRNA EST 24-JUL-2000
LOCUS BE418354
DEFINITION SCL031.D08R90531 ITEC SCL Wheat Leaf Library Triticum aestivum
cDNA clone SCL031.D08, mRNA sequence.
ACCESSION BE418354
VERSION BE418354.1 GI:9416200
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 988)
AUTHORS Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
Pecchlioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
Sorrells,M., Warburton,M. and Wenzel,G.
TITLE International Triticeae EST Cooperative (ITEC): Production of
International Triticeae EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticeae
JOURNAL Unpublished (2000)
COMMENT Contact: Cloutier S
Cereal Research Centre, Agriculture & Agri-Foods Canada
Winnipeg MT CANADA
Tel: 204 983 2340
Fax: 204 983 4604
Email: scloutier@em.agr.ca
International Triticeae EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.

FEATURES
source
1..988
/organism="Triticum aestivum"
/cultivar="Thatcher Lrl"
/db_xref="taxon:4565"
/clone="SCL031.D08"
/clone_lib="ITEC SCL Wheat Leaf Library"
/tissue_type="young leaf"
/dev_stage="after 24 hour challenge with LR-AVR1"
/notes="Vector: Lambda ZAP; 1.0 kbp average insert size."
BASE COUNT 227 a 199 c 244 g 247 t 71 others
ORIGIN

Query Match 17.8%; Score 494.4; DB 10; Length 988;
Best Local Similarity 96.9%; Pred. No. 1.1e-41;
Matches 564; Conservative 0; Mismatches 12; Indels 6; Gaps 6;

QY 1699 atggccattccagagctcatgaggagacgtgcaattgtctatctgtgattctgggat 1758
|||||
Db 7 AGGGCCATTCCAGAGCTCATGAGGAGGACGTGCAATTTGTCATGCTTGGATCTGGGAT 66
|||||

```

```

QY 1759 ccaatttttaagctogatgagatctaccagagtcaggttacaaggataaattccgtgga 1818
|||||
Db 67 CCAATTTTGAAGGCTGGATGATATCTACCGAGTCGAGTTACAAGGATAAATTCGCTGGA 126
|||||
QY 1819 tgggttgattagttccagtttccacagaaactcaggttcagttcgatattatttta 1878
|||||
Db 127 TGGTTCGGATTAGTGTTCAGTTTCCACAGAATAACTGCAGGTTCGATATATCTTA 186
|||||
QY 1879 atgccatcgagattgaaccttcggtcttaactaagctatatctatgcgaatatgtgaca 1938
|||||
Db 187 ATGCCATCGAGATTGAACCTTCGGTCTTAATCAGCTATATGCTATGCAATATGTTACA 246
|||||
QY 1939 -gttctctgtagtcatggaactgagggcctccagacacacagtcgaaccttcaaccttt 1997
|||||
Db 247 CGTTCCTGTAGTTCATGGAACCTGGGGCCTCCGAGACACAGTCGAGACCTTCAACCCCTT 306
|||||
QY 1998 tgggtcaaaaaggagaggggtacaggggtggtgtctctaccgcctaaacctggaacaagt 2057
|||||
Db 307 TGGTGTCAAAAAGGAGAGGGGTACAGGGTGGCGCTTCTCACCGCTAACCGTGGACAAGAT 366
|||||
QY 2058 qttgtgggcaattcgaaacccgcatgtcgacattcaggaggagacaaagccgtctctgggaggg 2117
|||||
Db 367 GTTGTGGGCATTCGGAACCGCATGTGACATTCAGGGGACACAAGCCGTTCTGGGAGGG 426
|||||
QY 2118 gctcatgaacggagggatgacgaagaccata-cgtgggaccatgcc-cgagacagtacg 2175
|||||
Db 427 GCTNATNACGGAGGATGACGAAGAACCATACCTGCTGGGACCATTGCCGCCGACGATCG 486
|||||
QY 2176 agcagatcttcgagtggtggtctcgtggaccaacctcactcgtca-tgtagacgggactggg 2234
|||||
Db 487 AGCAGATNTTTGAGTGGCGCTTCGTGGACCAACCTACNTCATGTTAGACNGGAGCTGGG 546
|||||
QY 2235 gag-gtccaaagtccagatctcc-ttcagctctgaagacatcc 2274
|||||
Db 547 GAGNGTCCAAAGTCGAGTCTCCTTTGAGCTTTGAANACATCC 588
|||||

```

```

RESULT 4
BG607242/c 557 bp mRNA EST 17-APR-2001
LOCUS WHE2493_B06_D1LZS Triticum monococcum early reproductive apex cDNA
DEFINITION library Triticum monococcum cDNA clone WHE2493_B06_D1L, mRNA
sequence.
ACCESSION BG607242
VERSION BG607242.1 GI:13657225
KEYWORDS EST.
SOURCE Triticum monococcum.
ORGANISM Triticum monococcum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 557)
AUTHORS Anderson,O.D., Chao,S., Dubcovsky,J., Echenique,V., Han,P.S., Hsia
,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L.,
Stamova,B. and Tong,J.C.
TITLE The structure and function of the expressed portion of the wheat
genomes - Early reproductive apex cDNA library from Triticum
monococcum
JOURNAL Unpublished (2001)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
FEATURES
source
1..557
/organism="Triticum monococcum"
/cultivar="DV92"

```



/db\_xref="taxon:4568"  
 /clone="WHE2493\_B06\_D11"  
 /clone\_lib="Triticum monococcum early reproductive apex  
 cDNA library"  
 /tissue\_type="Early reproductive apex"  
 /dev\_stage="Seven week-old plants"  
 /lab\_host="E. coli XL0UR"  
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
 Site\_1: EcoRI; Site\_2: XhoI; The tissue, total RNA, and  
 poly(A) RNA were prepared from apex at double-ridge stage  
 to terminal-spikelet stage during transition from  
 vegetative state to flower state, a cDNA library was made,  
 and the cDNA clones were in vivo excised at the  
 University of California, Davis (V. Echenique, B. Stamova  
 , J. Dubcovsky). Plasmid DNA preparations and DNA  
 sequencing were performed in the OD Anderson lab (all  
 other authors)."

BASE COUNT 140 a 161 c 118 g 138 t  
 ORIGIN

Query Match 14.6%; Score 403.8; DB 11; Length 557;  
 Best Local Similarity 88.9%; Pred. No. 2.1e-32;  
 Matches 472; Conservative 0; Mismatches 52; Indels 7; Gaps 3;  
 QY 2111 gggagggtctatgaagcgaggtgacgaaagaccatagctgggaccatgcc-ccgagc 2169  
 DB 557 GGGAGGGGCTCATGACGAGGCGATGACGAAAGACCATATGTGGACCATGCCCGCAGC 498  
 QY 2170 agtcagagcagattcttcagtggtggtctcttgacacacccctacgtcatgtagacggga 2229  
 DB 497 AGTACGACGAGATCTCGAATGGGCCCTCGTGGACCAACCCATATGTATGTAGACAGG 438  
 QY 2230 ctggggaggtccaaagtgcagctccctcagctctgaagacatctctcatcttcctcgc 2289  
 DB 437 CTGGGAGGTGCAAGCGCAGTCTCTTGAGCTCAGAAGACATGTTCTCATCTTCGCG 378  
 QY 2290 ggcgggaagatgataccctctgacattgcgttgcctgcctac-----agttagctgcgaat 2344  
 DB 377 GGCCTGGAAGGATACCCCTGTACATTCGCTGGAACCGGGCTCTCTCAATAGTGGCAAT 318  
 QY 2345 gcgcctctctgttctgctgcggttcgagaacatgatgacgctgctgctgcgcgcg 2404  
 DB 317 GCGCCTGCTTCG-TGGGTTCGCCGGTTCGAGAGTAAATGATGGCTGTGCTGCTGCGCGG 259  
 QY 2405 tgacagcttcgggtggacgacagttacagtttttggggaataagaagggatgctgctcag 2464  
 DB 258 TGACAGCTTCGGGTGATGACAGTTACAGTTTTCGGGAATAAGAGGGGTGCTGCAGG 199  
 QY 2465 gatggttaacgaaagaccacactcagatggcagcctctctgtccgtgttacagctgaaa 2524  
 DB 198 AATGTTAAACAGCAAGTTGCACTCAGCTGGCAGCCCTCTCGGTCCGTGTACAGCTGAAA 139  
 QY 2525 tcagaaacaaactgtaactctttagcttagcttagtgaattgtaagtttctgctctctgt 2584  
 DB 138 TCTGAACCAACTGGTGACTCTTTAGCCTTAGTGATGTGAAGTTTGTTCCTCTCTGT 79  
 QY 2585 atgttctgtctcttagctgacaaaatatattgacctgttggagaaatttat 2635  
 DB 78 ATGTTGTCTTCTCTAGCTGACAAATTTTGTGACCTGTGTGGAGAAATTTT 28

## RESULT 5

AW678087 562 bp mRNA EST 19-JUL-2000  
 LOCUS WS1\_13\_G10\_b1\_A002 Water-stressed 1 (WS1) Sorghum bicolor cDNA,  
 DEFINITION mRNA sequence.  
 ACCESSION AW678087  
 VERSION AW678087.1 GI:7551807  
 KEYWORDS EST.  
 SOURCE sorghum.  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 1 (bases 1 to 562)  
 Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt  
 ,L.H.  
 An EST database from Sorghum: water-stressed plants  
 Unpublished (2000)  
 Contact: Cordonnier-Pratt MM  
 Department of Botany  
 The University of Georgia  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 542 1805  
 Email: mmpratt@uga.edu  
 Sequences have been trimmed to exclude PolyA, vector and regions  
 below Phred quality 16. The threshold for highest quality sequence  
 is 20.  
 Seq primer: JEN REV  
 High quality sequence start: 17  
 High quality sequence stop: 553  
 POLYA=NO.

FEATURES  
 source

Location/Qualifiers  
 1..562  
 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4568"  
 /clone\_lib="Water-stressed 1 (WS1)"  
 /note="Organ: Mix of 5-week old plants on days 7 & 8 after  
 water was withheld; Vector: Lambda zap; Site\_1: XhoI;  
 Site\_2: EcoRI; The library was made from poly-A RNA in the  
 cloning vector lambda ZAP II. Clones to be sequenced were  
 prepared by mass excision."  
 BASE COUNT 153 a 126 c 142 g 140 t  
 ORIGIN

Query Match 14.2%; Score 393; DB 10; Length 562;  
 Best Local Similarity 84.1%; Pred. No. 2.5e-31;  
 Matches 466; Conservative 0; Mismatches 86; Indels 2; Gaps 2;  
 QY 1660 ggaagactggtaccagaaaggcattgattctctattataaaaggccattccagagctcatg 1719  
 DB 10 GNAAGATTGGATTATCAGAAA-GCATTGATCTCATTTCACTTATCATACACATCTCATG 68  
 QY 1720 agggagacgtgaatttgcgtctggtatctgggagatcccaatttttggaggctgagatg 1779  
 DB 69 CGGACGACCTTCAATTTGTTCATCTGGATCTGGTACCACAGAGCTCGAAGACTGGATG 128  
 QY 1780 agactaccagtcaggttacaggaataaattccgtgagtggttgatttagtcttcca 1839  
 DB 129 AGATCTACAGATCGGACTTCAGAGTAATAATTCGTGGATGGGTGGATTAGTGTCCA 188  
 QY 1840 gttcccccacagaataaactgcaggttgcgatatattgttaatgccaatcgagattgaaacct 1899  
 DB 189 GTTTCACCGAATAACTGCCGGCTGGATATATTGTTAATGCCATCCAGATTCGAACCT 248  
 QY 1900 tgcggtcttaacagctatgtatgtatgcataatggtacagttcctgtagttcaggaact 1959  
 DB 249 TGTGGTCTCAATCAGCTATATGCTATGCAGTATGGCACAGTTCCTGTGTGCCATGCAACT 308  
 QY 1960 gggggcctccgagacacagtcgagaccttcaaccttttgggtcaaaagagagagaggggt 2019  
 DB 309 GGGGCCCTTAGAGATACTGTGGAGAACTTCAACCCCTTCGGTGAGATGGGAACAGGGT 368  
 QY 2020 acaggggtggcggtctctaccgctaaaccttggaagaagtgtgtgtgggcatctgcgaaccgcg 2079  
 DB 369 ACAGGGTGGGCATTCGACCCCTTAACACGGAACAACATGTTGTGAGATTCGGAACATGCA 428  
 QY 2080 atgtcagattcagggagacacagccctcctggagggtcctcatgagcagcagcagcagc 2139  
 DB 429 ATTTGACATACAAAGAGGACAAATCTTCCTGGGAAGGGCTCATGAAGAGGGGCTGTCA 488  
 QY 2140 aaagaccatactggtggaccatgcc-ccgagcagtcacgagcagatcttcgagtggtggccttc 2198



Db 489 AAGACTTCACGTGGGACCATCCGCTGAACAATAACGAACAACTTCCAGTGGCCTTC 548

QY 2199 gtggaccacacctta 2212

Db 549 ATCGATCGACCTTA 562

RESULT 6

AV832689 454 bp mRNA EST 22-JUN-2001

LOCUS AV832689 K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare top three leaves adult, heading stage Hordeum vulgare subsp. vulgare cDNA clone baal2b23, mRNA sequence.

DEFINITION

ACCESSION AV832689.1 GI:14524778

VERSION AV832689

KEYWORDS EST

SOURCE Hordeum vulgare subsp. vulgare.

ORGANISM Hordeum vulgare subsp. vulgare

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.

AUTHORS Sato, K.

TITLE Barley EST sequencing project in NIG and Okayama Univ

JOURNAL Unpublished (2001)

COMMENT Contact: Kazuhiro Sato  
Research Institute for Bioreources  
Okayama University, Barley Germplasm Center  
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan  
Email: kazzato@rib.okayama-u.ac.jp  
URL: http://www.rib.okayama-u.ac.jp/barley/  
Sato, K., Saisho, D., Takeda, K., Shini, T. and Kohara, Y. Direct submission;  
database: http://www.shigen.nig.ac.jp/barley/Barley.html.

FEATURES

source

1. .454

Location/Qualifiers

/organism="Hordeum vulgare subsp. vulgare"

/cultivar="Haruna Nijo"

/db\_xref="taxon:112509"

/clone="baal2b23"

/clone\_lib="K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare top three leaves adult, heading stage"

/tissue\_type="top three leaves"

/dev\_stage="adult, heading stage"

BASE COUNT 112 a 84 c 126 g 130 t 2 others

ORIGIN

Query Match 14.0%; Score 388; DB 10; Length 454;

Best Local Similarity 95.1%; Pred. No. 8.8e-31;

Matches 411; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 592 gaaggattgctgagattccatcgacagcataattgtgctgcaagtgcagcagattct 651

Db 24 GGAGGATTGCTGAGGATTCC-TCGACACCATAGTTGTGGCTGCAAGTCAGCAGGATTCT 82

QY 652 gagatcatgatgcaatgacaaacctcaagctaaagtacacgtacgtatggtgtgtg 711

Db 83 GAGATCATGATGCCAACGACCACTCTAGCTAAAGTTACACGATGATGTTGTGTG 142

QY 712 actggtgaagctgctctcttatcaaaagtcagggggttggagatgttgggttcgta 771

Db 143 ACTGTTGAAGCTGCTCTTATCAAGTCAGGGGGCTGGGAGATGTTTGTGTTGCTGTG 202

QY 772 ccaattgcttctgctgctggtggtcaccagtgatggtgttaatgccaaataactaaat 831

Db 203 CCAATTGCTCTTGTCTGCTGGTGCACCGTGTGATGTTGTTAAATGCCGAGATACTTAAT 262

QY 832 gggctctctgaataaaactatgcaaaagcattatcacactgcgaagcacattcaagattcca 891

Db 263 GGGACCTCTGATAAAACTATGCAAAAGGCATTATACACCGGGGAAGCACATTAAGATTCCA 322

QY 892 tgctttggggatcacatgaagtgaaccttttttcatagatagagacaacgctcgattgg 951

Db 323 TGCTTCGGGGATCACATGAAGTCACCTTTTTCATGAGTATAGACAACGTCGATGG 382

QY 952 gtgtttgcatcatcctcgatcatatcacagaccaggaagtttatatgagataattttggt 1011

Db 383 GTGTTCGTTGATCATCATATCATAGACCAGGAAGTTTGTATGGAGATAATTTTGGT 442

QY 1012 gcttttggtgat 1023

Db 443 GCITTTGGTGAT 454

RESULT 7

BF482660 408 bp mRNA EST 06-DEC-2000

LOCUS WHE2301-2304\_E13\_E13S Wheat pre-anthesis spike cDNA library

DEFINITION Triticum aestivum cDNA clone WHE2301-2304\_E13\_E13, mRNA sequence.

ACCESSION BF482660

VERSION BF482660.1 GI:11565884

KEYWORDS EST

SOURCE bread wheat.

ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.

AUTHORS Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Seaton, C.L. and Tong, J.C.

TITLE The structure and function of the expressed portion of the wheat genomes - Pre-anthesis spike cDNA library

JOURNAL Unpublished (2000)

COMMENT Contact: Olin Anderson  
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA  
Tel: 5105595773  
Fax: 5105595818  
Email: oanderson@pw.usda.gov  
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20  
Seq primer: StrataGene SK primer.

FEATURES

Location/Qualifiers

1. 408

/organism="Triticum aestivum"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="WHE2301-2304\_E13\_E13"

/clone\_lib="Wheat pre-anthesis spike cDNA library"

/tissue\_type="Spike before anthesis"

/dev\_stage="Adult plant"

/lab\_host="E. coli SOLR"

/note="vector: Lambda Uni-ZAP XR, excised phagemid; Site 1: EcoRI; Site 2: XhoI; plants were grown in the greenhouse. Whole spike with awns trimmed, white, and yellow anther were collected and total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 108 a 76 c 99 g 125 t

ORIGIN

Query Match 13.7%; Score 380.8; DB 11; Length 408;

Best Local Similarity 95.8%; Pred. No. 4.9e-30;

Matches 391; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1558 tattctgtcgtgacctctctggaaggccaaatgtaagctgaattgcagaaggagttg 1617



Db	1	TATTCTGTCGATGACCTCTCTGGAAGGCCAAATGTAAGCTGAATTGCACAGGAGCTG	60
QY	1618	ggtttacctgtaaggagatgtctctgattggtctttatttggaagactggattaccag	1677
Db	61	GGTTTTACCTGTAAGGAGAGATGTTCTCTGATGGCTTATTGGAAGACTGGATTACCAC	120
QY	1678	aaaggcattgatctcattaaatggccattccagagctcatcagggagagactgcaattt	1737
Db	121	AAAGGCAATGATCTCATTAATTAATGGCCATTCAGAACTCATGAGGAGGACGTACACTTT	180
QY	1738	gtatcgttgatctggggatccaatttttgaagcgtgatgatctaccagactcgagt	1797
Db	181	GTATGCTTGGATCTGGGGATCCAAATTTTGAAGGCTGGATGAGATCTACCGAGTCTAGT	240
QY	1798	tacaagataaaatccgtggtggtggttagttagtctcagtttccacacgaataact	1857
Db	241	TACAAGGATAAATTCGTGGATGGTGGATTTACTGTCCCAATTTCCACACAAATACT	300
QY	1858	gcaggttcgatataattgtaagccatcgagatttgaaccttcgcttatacagcta	1917
Db	301	GCATGTTGCCGATATATTGTTAATGCCATCCATGTTTGAACCTTGTGTCITTAATCAACTA	360
QY	1918	tatcgtgcaaatggtacagttccctgtagttcattcagtggaactgggggc	1965
Db	361	TATGCTATGCAATATGTACAGTTCTCTGTAGTTTCATGGAATGGGGGC	408
RESULT	8		
LOCUS	BF113316	596 bp mRNA	EST 18-MAY-2001
DEFINITION	EST440906 tomato breaker fruit Lycopersicon esculentum cDNA clone		
ACCESSION	BF113316		
VERSION	BF113316.1	GI:10943006	
KEYWORDS	EST.		
SOURCE	tomato.		
ORGANISM	Lycopersicon esculentum		
REFERENCE	1 (bases 1 to 596)		
AUTHORS	Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley S.D.		
TITLE	Generation of ESTs from tomato fruit tissue, breaker stage		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: CUGI Clemson University Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> .		
FEATURES	Location/Qualifiers		
source	1..596		
	/organism="Lycopersicon esculentum"		
	/cultivar="TA496"		
	/db_xref="taxon:4081"		
	/clone.lib="cLEG44E24"		
	/clone.lib="tomato breaker fruit"		
	/tissue_type="pericarp"		
	/dev_stage="breaker"		
	/lab_host="SOLR"		
	/note="Vector: pBluescriptSKmQuadapt; Site_1: EcoRI; Site_2: XhoI; Fruit were harvested at the breaker stage (first sign of lycopen accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."		
BASE COUNT	167 a 120 c 149 g 160 t		
ORIGIN			
Query Match	11.8%; Score 325.6; DB 11; Length 596;		

Best Local Similarity	71.6%;	Pred. No. 1.6e-24;	
Matches	427;	Conservative	0; Mismatches 169; Indels 0; Gaps 0;
QY	1156	cttcttgctgcacaaatagaccatacagtggtttacagagattcccgagacccttatt	1215
Db	1	CTTTTAGCAGCCAAAGTATCGTCTTATGTTTAAAGGATGCTCGTAGTATGTCGCA	60
QY	1216	atacataatttagcacatcaggtgtggagcctgcaagtcacatataccttgatctgggattg	1275
Db	61	ATACACACATTTGCACATTCAGGAGCTGGAGCCTGCAGCAACCTACAATAATTTGGGATTG	120
QY	1276	cctctcgaatggtatggagcttttagaaggggtattccagaatgggcaaggagccatcc	1335
Db	121	CCTCCTCAATGCTATGGAGCACTTGAATGATATTTCCACATGGCAAGGGCCCATCGG	180
QY	1336	cttgacaagggtgagcgacttaactttttgaaaggcgaggttgtgacagcagatcggaatt	1395
Db	181	CTTGACACTGGGTGAACACAGTGAATGTTTCAAGGGGCAATCTCAGTTGCTGATCGGATA	240
QY	1396	gtgacctcagtcaggtttattcattcagggggtcacaactcgtcgaagggtggacagggcctc	1455
Db	241	CTGACAGTTAGCCAGGGATCTCATGGGAAATAAACAACTCCTGAAGGGGATATGGGTTA	300
QY	1456	aatgagctcttaagctcccccgaagaaagtgtattgaatgggaattgtaaatggaattgacatt	1515
Db	301	CATGAGCTGCTGAGCAGTAGACAGTCAGTCTTAAATGGAATTACTAATGGAATAGATGTT	360
QY	1516	aatgattggaaccccccacacagacaagtgctctccctcatcattattctctgatgacctc	1575
Db	361	AATGATTGGAAACCCGCGACAGATGAGCATATATGCTTCGATTAATCTCCATCAATGACCTC	420
QY	1576	tctggaagggccaaaagttaaagctgaatgcagaaaggagttgggtttactctgaaggag	1635
Db	421	TCTGGAAGGCTCAGTGCAGACTGATCTGCAAAAGAACTTGGGCCCTTCCAATTCACCT	480
QY	1636	gattctcctctgattggtcttattggaagactggtattaccagaaagcatttgatctccatt	1695
Db	481	GATTGTCGGCTGATGGATTTATTGGAGGCTGGACTACCCAGAAAGGTGTTGACATAATC	540
QY	1696	aaatggccattccagagctcatcaggggagcgtgcaattttcattctgattgattc	1751
Db	541	CTGTCAGCAATTCAGAACTTTTTCAGAGAGGATGTCCTCAATTTGTAATGCTTGGATC	596
RESULT	9		
LOCUS	BE434991	588 bp mRNA	EST 18-MAY-2001
DEFINITION	EST406069 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA		
ACCESSION	BE434991		
VERSION	BE434991.1	GI:9432834	
KEYWORDS	EST.		
SOURCE	tomato.		
ORGANISM	Lycopersicon esculentum		
REFERENCE	1 (bases 1 to 588)		
AUTHORS	Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley S.D.		
TITLE	Generation of ESTs from tomato fruit tissue, breaker stage		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: CUGI Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> 5 prime sequence.		
FEATURES	Location/Qualifiers		
source	1..588		







REFERENCE 1 (bases 1 to 401)  
AUTHORS Cordonnier-Pratt M.M., Gingle A., Sudman M., Marsala C. and Pratt L.H.  
TITLE An EST database from Sorghum: floral-induced meristems  
JOURNAL Unpublished (2000)  
COMMENT Department of Botany  
The University of Georgia  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 542 1805  
Email: mmpratt@uga.edu  
Sequences have been trimmed to exclude polyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.  
Seq primer: JEN REV  
High quality sequence stop: 361  
POLYA-No. Location/Qualifiers  
1. .401

FEATURES  
source  
1. .401  
/organism="Sorghum propinquum"  
/db\_xref="taxon:132711"  
/clone\_lib="Floral-Induced Meristem 1 (FMI)"  
/note="Organ: Floral-Induced Meristems; Vector: pBlueScript II from Lambda Zap II; Site\_1: XhoI; Site\_2: EcoRI; mature plants were placed in a growth chamber for 15 days with 16 hr darkness and 8 hr light (flowering is induced by short-day conditions); 16 days after being returned to the greenhouse under natural long days during late April/early May, meristems were harvested. The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."  
102 a 72 c 104 g 123 t

BASE COUNT 102 a 72 c 104 g 123 t  
ORIGIN  
Query Match 10.0%; Score 276; DB 11; Length 401;  
Best Local Similarity 90.7%; Pred. No. 1.9e-19;  
Matches 294; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
Qy 1262 ctgactgggattgctcctcctgagtgatggagcttttagaattgggtatttccagaaatggg 1321  
Db 1 CTGACCTTGGTGGCCACCTGAATGGAGCTGGAGCTCTGGAGTGGGTATTCCTGAATGGG 60  
Qy 1322 caaggaggatgcctcctgacagggtgaggcaggttaacttttgaaggagcagttgtga 1381  
Db 61 CAAGGAGGCATGCCCTTGACAGGGGTGAGGCGATTAATTTTGAAGGTGCGAGTTGCA 120  
Qy 1382 cagcagatcggttgaccgacagtgagggttattcattggaggagtcacaaactgctgaag 1441  
Db 121 CAGCAGATCGAATTTGACTGTCACTAGAGGTATTTCATGGGAGTCAACACTGCTGAAG 180  
Qy 1442 gtggacagggcctcaatgagctcttaagctcccgaaaagtgtatgaaatggaaatgaa 1501  
Db 181 GTGGACAGGGCCTCAATGAGCTTTAAGCTCCGAAAGAGTGTATTAACGGAATGTAA 240  
Qy 1502 atggaattgacattaatgattggaacccaccacagacaagtgctcctcattcattatt 1561  
Db 241 ATGGGAATTCACATTAATGATTTGGAACCTTGGCAGCGGACATATGTATCCCTGCTCATTA 300  
Qy 1562 ctgtcgatgacacctctctggaagg 1585  
Db 301 CTCTTGATGACCTCTCTGGAAG 324

RESULT 12  
AW932199 564 bp mRNA EST 18-MAY-2001  
LOCUS EST358042 tomato fruit mature green, TAMU Lycopersicon esculentum  
DEFINITION CDNA clone cLEF47J7 5', mRNA sequence.

AW932199  
AW932199.1 GI:8107600  
EST  
tomato.  
SOURCE Lycopersicon esculentum  
ORGANISM  
REFERENCE 1 (bases 1 to 564)  
AUTHORS Alcalá, J., Vrebalov, J., White, R., Matern, A.L., Holt, I.E., Liang, F., Upson, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S., Ronning, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.  
COMMENT Generation of ESTs from tomato fruit tissue  
Unpublished (1999)  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: http://www.genome.clemson.edu/orders/index.html  
5 prime sequence.  
Location/Qualifiers  
1. .564  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone\_lib="cLEF47J7"  
/clone\_lib="tomato fruit mature green, TAMU"  
/tissue\_type="fruit pericarp"  
/dev\_stage="mature green (3-5 days pre-ripening)"  
/lab\_host="SOLR"  
/note="Vector: pBlueScript SK(-); Site\_1: EcoRI; Site\_2: XhoI; cLEF - Fruit were tagged at the lcm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp."  
150 a 115 c 140 g 159 t

FEATURES  
source  
1. .564  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone\_lib="cLEF47J7"  
/clone\_lib="tomato fruit mature green, TAMU"  
/tissue\_type="fruit pericarp"  
/dev\_stage="mature green (3-5 days pre-ripening)"  
/lab\_host="SOLR"  
/note="Vector: pBlueScript SK(-); Site\_1: EcoRI; Site\_2: XhoI; cLEF - Fruit were tagged at the lcm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp."  
150 a 115 c 140 g 159 t  
ORIGIN

Query Match 9.0%; Score 250.6; DB 10; Length 564;  
Best Local Similarity 67.7%; Pred. No. 6e-17;  
Matches 381; Conservative 0; Mismatches 154; Indels 28; Gaps 1;  
Qy 1136 atgcacgcttgctgcagctctcttctgcacaaatagaccatcggtgtttacagag 1195  
Db 1 ATGCTTCCCTGGTTCCTTTACTTTTACGACGCCAAGTATCGTCTTATGGTGTTCACAGG 60  
Qy 1196 attcccgagcacccctgttatataatattagcacatc----- 1234  
Db 61 ATGCTCTAGTATTGTCGCAATACACAACATTGCACATCAGATCGTCTCTTGTGTGATAT 120  
Qy 1235 -----agggtgtggagcctgcgaagtacatatctgactctgggattgctcctcgaatgg 1287  
Db 121 TTTCATATAGGAGTGGAGCTGCGACCACTACAAATAATTTGGGATTTGCCCTCCCTCAATGG 180  
Qy 1288 tatggagctttagaatgggtatttcccaagaatggcagaggagcatgcccttgacaagggt 1347  
Db 181 TATGGAGCAGTGAATGGATATTTCCTCACATGGGCAAGGCCCATGGCTTGACACTGGT 240  
Qy 1348 gaggcagtaacttttgaaggagcaggtgtgacagcagatcgaggttgagccgtcagt 1407  
Db 241 GAACAGTGAATGTTTTGAAGGGGCAATCTCAGTTGCTGATCGGATACGACAGTTAGC 300  
Qy 1408 cagggttattcattggagggttcacaactgctgaaggtggagagccctcaatgaactctta 1467  
Db 301 CAGGGATACTCATGGAAATAACAACCTCCTGAAGGGGATATGGCTACATGAGCTGCTG 360  
Qy 1468 agctcccccacaaagtgtattgaatgggaattgtaaatgaattgacattgaattgggaac 1527  
Db 361 AGCAGTAGACAGTCTGCTTCTTAATGGAATTAATAAGTAAGATGTTAATGATTGGAAC 420



Qy 1528 ccacacacagacaagtgtctcctcatcattatctgctgatgacctctctggaagcc 1597  
Db 421 CCCTGCACAGATGAGCATATTGCTTCGATTACTIONCAATGACCCTCTGGAAAGCT 480  
Qy 1588 aaatgtaagctgaattgcagaagtggtggtttacctgtaaaggagtagtccctctg 1647  
Db 481 CAGTCAAGACTGTACTGCAAAAGAAGCTGGCGCTCCCAATCACCCTGATTGTCGCTG 540  
Qy 1648 attggcttattggaagactgga 1670  
Db 541 ATTGGATTATTGGAAGGCTGGA 563

AW759569

### DEFINITION

VERSION

**SOURCE**  
**ORGANIZATION**

## AUTHORS

**TITLE**

**COMMENT**

source







Search completed: March 28, 2002, 13:48:18  
Job time: 5992 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2002, 15:16:00 ; Search time 238.29 Seconds  
(without alignments)  
880.896 Million cell updates/sec

Title: US-09-674-824-2

Perfect score: 4044

Sequence: 1 MAATGVGAGCLAFPSVRLRAD.....SDGSLSVRTAEIRNQLVTL 756

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending Patents\_AA\_Main.\*
- 1: /cgn2\_6/ptodata/1/paa/PTCUS\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*
  - 8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*
  - 13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*
  - 17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*
  - 18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*
  - 19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*
  - 20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*
  - 21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*
  - 22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*
  - 23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*
  - 24: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2557.5	63.2	583	18	US-09-402-254-53
2	2557.5	63.2	583	20	US-09-625-406-13
3	2390.5	59.1	539	20	US-09-625-406-21
4	2189	54.1	435	24	US-60-312-544-9109
5	1992	49.3	641	20	US-09-606-304-10
6	1718.5	42.5	459	20	US-09-606-304-4
7	1374	34.0	303	24	US-60-324-109-29423
8	1337	33.1	297	24	US-60-312-544-5760
9	1276	31.6	535	24	US-60-324-109-21228

10	1225.5	30.3	670	18	US-09-402-254-51	Sequence 51, Appl
11	1215	30.0	767	20	US-09-606-304-8	Sequence 8, Appl
12	1209.5	29.9	804	18	US-09-402-254-49	Sequence 49, Appl
13	1205.5	29.8	698	20	US-09-625-406-11	Sequence 11, Appl
14	1196.5	29.6	801	17	US-09-388-743-26	Sequence 26, Appl
15	1184	29.3	558	20	US-09-606-304-6	Sequence 6, Appl
16	1142.5	28.3	690	17	US-09-388-743-6	Sequence 6, Appl
17	1068	26.4	669	20	US-09-625-406-9	Sequence 9, Appl
18	1055.5	26.1	440	24	US-60-324-109-17143	Sequence 17143, A
19	973.5	24.1	385	24	US-60-312-544-10029	Sequence 10029, A
20	891.5	22.0	636	24	US-60-288-315-4	Sequence 4, Appl
21	881	21.8	637	18	US-09-402-254-47	Sequence 47, Appl
22	880	21.8	477	19	US-09-583-110-3812	Sequence 3812, Ap
23	880	21.7	484	15	US-09-107-433-4468	Sequence 4468, Ap
24	877	21.7	609	20	US-09-625-406-7	Sequence 7, Appl
25	875.5	21.6	616	17	US-09-388-743-14	Sequence 14, Appl
26	870	21.5	604	21	US-09-731-166-4	Sequence 4, Appl
27	865	21.4	533	20	US-09-625-406-5	Sequence 5, Appl
28	860	21.3	614	17	US-09-388-743-18	Sequence 18, Appl
29	852.5	21.1	600	17	US-09-388-743-22	Sequence 22, Appl
30	840	20.8	342	24	US-60-312-544-10086	Sequence 10086, A
31	828.5	20.5	615	17	US-09-388-743-2	Sequence 2, Appl
32	826	20.5	599	24	US-60-288-315-5	Sequence 5, Appl
33	824	20.4	609	24	US-60-324-109-18806	Sequence 18806, A
34	821	20.3	609	24	US-60-288-315-2	Sequence 2, Appl
35	796.5	19.7	479	23	US-09-902-540-14647	Sequence 14647, A
36	696	17.2	480	21	US-09-739-449-8399	Sequence 8399, Ap
37	696	17.2	480	22	US-09-803-110-8399	Sequence 8399, Ap
38	691.5	17.1	548	16	US-09-252-691-8864	Sequence 8864, Ap
39	691.5	17.1	548	16	US-09-252-691-8864	Sequence 8864, Ap
40	681.5	16.9	511	18	US-09-489-039A-11033	Sequence 11033, A
41	678.5	16.8	477	3	US-07-731-226-2	Sequence 2, Appl
42	678.5	16.8	477	4	US-08-016-881A-12	Sequence 12, Appl
43	678.5	16.8	477	8	US-08-484-434A-12	Sequence 12, Appl
44	678.5	16.8	477	17	US-09-384-361-12	Sequence 12, Appl
45	650	16.1	236	24	US-60-324-109-25214	Sequence 25214, A

ALIGNMENTS

RESULT 1  
US-09-402-254-53  
; Sequence 53, Application US/09402254  
; GENERAL INFORMATION:  
; APPLICANT: Guan, Hanping  
; TITLE OF INVENTION: PLANT LIKE STARCHES AND THE METHOD OF MAKING THEM IN  
; FILE OF INVENTION: HOSTS  
; FILE REFERENCE: 2461-52  
; CURRENT APPLICATION NUMBER: US/09/402,254  
; CURRENT FILING DATE: 1999-10-01  
; EARLIER APPLICATION NUMBER: PCT/US98/06660  
; EARLIER FILING DATE: 1998-04-03  
; EARLIER APPLICATION NUMBER: 60/042,939  
; EARLIER FILING DATE: 1997-04-04  
; NUMBER OF SEQ ID NOS: 77  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 53  
; LENGTH: 583  
; TYPE: PRT  
; ORGANISM: Zea mays  
; US-09-402-254-53

Query Match 63.2%; Score 2557.5; DB 18; Length 583;  
Best Local Similarity 86.7%; Pred. No. 1.5e-213;  
Matches 476; Conservative 25; Mismatches 45; Indels 3; Gaps 2;  
QY 45 VAEISREGPAARPAQOOOLAPPLVPGFLAPPPAPAPASPAOTPPPLPDAGVGLAPDLLL 104  
DB 2 VAEISREGPAARPPALLAPPLVPGFLA-PPAEPTGEPASTPPPPVDPAGLGLDGL--LEP 58



QY 105 EGAEDSIDSIIVASEQDSEIMDANEQOAKVTRISIVFTGEAAYAKSGGLGDCVCSL 164  
DB 59 EGAEGSIDNTVVASEQDSEIVVGEQAKARAKVTQSVFVTGEASPYAKSGGLGDCVCSL 118  
QY 165 PIALAARGHRVVMVMPRYLNGSDKNYAKALYAKIKIPCFGGSHEVTFHEYRDNDV 224  
DB 119 PVALAARGHRVVMVMPRYLNGTSDKNYANAFYTEKHIRIPCFGGEHEVTFHEYRDSVD 178  
QY 225 VFVDHPSYHRPGSLYGNFAGFQDNFRYTLICYAACEAPLIILELGGYIYGQNCMEVND 284  
DB 179 VFVDHPSYHRPGSLYGNFAGFQDNFRYTLICYAACEAPLIILELGGYIYGQNCMEVND 238  
QY 285 WHASLVVLLAAKRYRPGYVVRDSTRILVTHNLAHQGVPEASTYVDLGLPPEWYGALEWVF 344  
DB 239 WHASLVVLLAAKRYRPGYVVRDSTRILVTHNLAHQGVPEASTYVDLGLPPEWYGALEWVF 298  
QY 345 PEWARRHALDKGEAVNFKGAVVTADRIVTVSQGSYSEVTTAEGGQGLNELLSSRKSVLN 404  
DB 299 PEWARRHALDKGEAVNFKGAVVTADRIVTVSQGSYSEVTTAEGGQGLNELLSSRKSVLN 358  
QY 405 GIVNGIDINDWNPATDKCLPHHYSVDDLSGKAKCKALQKELGLPIRPDVLGFTGRLD 464  
DB 359 GIVNGIDINDWNPATDKCIPCHYSVDDLSGKAKCKALQKELGLPIRPDVLGFTGRLD 418  
QY 465 YQKGDILKMAIPELMREDVQFVNLGSDPIFEGWMRSTESSYKDKFRGWGFSVPVSHR 524  
DB 419 YQKGDILQIILPDLMRDQFVNLGSDPIFEGWMRSTESSYKDKFRGWGFSVPVSHR 478  
QY 525 ITAGCDIILMPSRPEPCGLNQLYAMQYGTVPVHGTGGLRDTVETFPNFGAKGEEGTGWA 584  
DB 479 ITAGCDIILMPSRPEPCGLNQLYAMQYGTVPVHGTGGLRDTVETFPNFGGEGQGTGWA 538  
QY 585 FSPLTVDKM 593  
DB 539 FAPLTENM 547

RESULT 2  
US-09-625-406-13  
; Sequence 13, Application US/09625406  
; GENERAL INFORMATION:  
; APPLICANT: Keeling, Peter  
; APPLICANT: Guan, Hanping  
; TITLE OF INVENTION: Starch Encapsulation  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/625,406  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/941,445  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winner, Ellen P.  
; REGISTRATION NUMBER: 28,547  
; REFERENCE/DOCKET NUMBER: 89-97  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 583 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-625-406-13  
  
Query Match 63.2%; Score 2557.5; DB 20; Length 583;  
Best Local Similarity 86.7%; Pred. NO. 1.5e-213;  
Matches 476; Conservative 25; Mismatches 45; Indels 3; Gaps 2;  
  
QY 45 VAELESREGPAARPAQOQQLAPPLVGFGLAPPAPPAQSPAPTPQPLPDAGVCELAPDLLL 104  
DB 2 VAELESREGPAARPLPALLAPPLVGFGLA-PPAETGEPASTPPPPPDAGLGDLG--LEP 58  
QY 105 EGAEDSIDSIIVASEQDSEIMDANEQOAKVTRISIVFTGEAAYAKSGGLGDCVCSL 164  
DB 59 EGAEGSIDNTVVASEQDSEIVVGEQAKARAKVTQSVFVTGEASPYAKSGGLGDCVCSL 118  
QY 165 PIALAARGHRVVMVMPRYLNGSDKNYAKALYAKIKIPCFGGSHEVTFHEYRDNDV 224  
DB 119 PVALAARGHRVVMVMPRYLNGTSDKNYANAFYTEKHIRIPCFGGEHEVTFHEYRDSVD 178  
QY 225 VFVDHPSYHRPGSLYGNFAGFQDNFRYTLICYAACEAPLIILELGGYIYGQNCMEVND 284  
DB 179 VFVDHPSYHRPGSLYGNFAGFQDNFRYTLICYAACEAPLIILELGGYIYGQNCMEVND 238  
QY 285 WHASLVVLLAAKRYRPGYVVRDSTRILVTHNLAHQGVPEASTYVDLGLPPEWYGALEWVF 344  
DB 239 WHASLVVLLAAKRYRPGYVVRDSTRILVTHNLAHQGVPEASTYVDLGLPPEWYGALEWVF 298  
QY 345 PEWARRHALDKGEAVNFKGAVVTADRIVTVSQGSYSEVTTAEGGQGLNELLSSRKSVLN 404  
DB 299 PEWARRHALDKGEAVNFKGAVVTADRIVTVSQGSYSEVTTAEGGQGLNELLSSRKSVLN 358  
QY 405 GIVNGIDINDWNPATDKCLPHHYSVDDLSGKAKCKALQKELGLPIRPDVLGFTGRLD 464  
DB 359 GIVNGIDINDWNPATDKCIPCHYSVDDLSGKAKCKALQKELGLPIRPDVLGFTGRLD 418  
QY 465 YQKGDILKMAIPELMREDVQFVNLGSDPIFEGWMRSTESSYKDKFRGWGFSVPVSHR 524  
DB 419 YQKGDILQIILPDLMRDQFVNLGSDPIFEGWMRSTESSYKDKFRGWGFSVPVSHR 478  
QY 525 ITAGCDIILMPSRPEPCGLNQLYAMQYGTVPVHGTGGLRDTVETFPNFGAKGEEGTGWA 584  
DB 479 ITAGCDIILMPSRPEPCGLNQLYAMQYGTVPVHGTGGLRDTVETFPNFGGEGQGTGWA 538  
QY 585 FSPLTVDKM 593  
DB 539 FAPLTENM 547

RESULT 3  
US-09-625-406-21  
; Sequence 21, Application US/09625406  
; GENERAL INFORMATION:  
; APPLICANT: Keeling, Peter  
; APPLICANT: Guan, Hanping  
; TITLE OF INVENTION: Starch Encapsulation  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:



APPLICATION NUMBER: US/09/625,406  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/941,445  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Winner, Ellen P  
REGISTRATION NUMBER: 28,547  
REFERENCE/DOCKET NUMBER: 89-97  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 539 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-625-406-21

Query Match 59.1%; Score 2390.5; DB 20; Length 539;  
Best Local Similarity 88.4%; Pred. No. 4.9e-199;  
Matches 444; Conservative 23; Mismatches 32; Indels 3; Gaps 2;  
QY 95 VGELA-PDLLL--EGIAEDSIDSIIVAAEQDSEIMDANQPOAKVTRISIVFTVGEAAPY 151  
Db 2 VAELSRDLGLEPGLAEGSIDNTVVVASQDSEIVGVKEQAKVTSIVFTVGEASPY 61  
QY 152 AKSGGLDGVCSLPALAAARGHRVWVMPRYLNGSSDKNYAKALYTAHKKIKPCFGGSHE 211  
Db 62 AKSGGLDGVCSLPALAAARGHRVWVMPRYLNGTSDKNYANAFYTEKHIRIPCFGGEHE 121  
QY 212 VTFEHEYRDVWVDFVDPHPSYHRPGSLYGNFCAFQDNQFRYTLICYAACEAPLILELGG 271  
Db 122 VTFEHEYRDVWVDFVDPHPSYHRPGSLYGNFCAFQDNQFRYTLICYAACEAPLILELGG 181  
QY 272 YIYGONCMFYVNDWHASLVPLVLAARYPYGVYRDSRSTLVIHNLAHQGVPEASTYPDLG 331  
Db 182 YIYGONCMFYVNDWHASLVPLVLAARYPYGVYRDSRSTLVIHNLAHQGVPEASTYPDLG 241  
QY 332 LPPEWYGALEWVPEWARRHALDKGEAVNFKAGVAVTADRIYTVSGYSGWVTTABGGOG 391  
Db 242 LPPEWYGALEWVPEWARRHALDKGEAVNFKAGVAVTADRIYTVSGYSGWVTTABGGOG 301  
QY 392 LNELLSRKSVLNGIYVNGIDINDWNPATDKCIPCHYSVDLGSKAKCAELKELGLPVR 451  
Db 302 LNELLSRKSVLNGIYVNGIDINDWNPATDKCIPCHYSVDLGSKAKCAELKELGLPVR 361  
QY 452 EDVPLIGFGRLDYQKIDILKMAIPELMREDVQFVNLGSGDPIFEGWMRSTESSYKDKF 511  
Db 362 PDVPLIGFGRLDYQKIDILKMAIPELMREDVQFVNLGSGDPIFEGWMRSTESSYKDKF 421  
QY 512 RGVWGFSPVSHRITAGCDILLMPSEFPCGLNOLYAMQYGVVHVHGTGGLRDVTFENFN 571  
Db 422 RGVWGFSPVSHRITAGCDILLMPSEFPCGLNOLYAMQYGVVHVHGTGGLRDVTFENFN 481  
QY 572 PEGAKEEGTGNAFSLPTVDMK 593  
Db 482 PEGAKEEGTGNAFSLPTVDMK 503

RESULT 4  
US-60-312-544-9109  
Sequence 9109, Application US/60312544  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Edgerton, Michael D  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jingdong  
APPLICANT: Stein, Joshua

TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT  
FILE REFERENCE: 38-10(52726)A  
CURRENT APPLICATION NUMBER: US/60/312,544  
CURRENT FILING DATE: 2001-08-15  
NUMBER OF SEQ ID NOS: 10730  
SEQ ID NO 9109  
LENGTH: 435  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3150-035-F10\_FLI  
US-60-312-544-9109

Query Match 54.1%; Score 2189; DB 24; Length 435;  
Best Local Similarity 91.3%; Pred. No. 1.4e-181;  
Matches 397; Conservative 16; Mismatches 22; Indels 0; Gaps 0;  
QY 159 DVCGLPPIALAAARGHRVWVMPRYLNGSSDKNYAKALYTAHKKIKPCFGGSHEVTFHFHEY 218  
Db 1 DVCGLPPIALAAARGHRVWVMPRYLNGTSDKNYANAFYTEKHIRIPCFGGEHEVTFHFHEY 60  
QY 219 RONVDWVFDHPSYHRPGSLYGNFCAFQDNQFRYTLICYAACEAPLILELGGYIYGNC 278  
Db 61 RDSVDWVFDHPSYHRPGSLYGNFCAFQDNQFRYTLICYAACEAPLILELGGYIYGNC 120  
QY 279 MFVVDWHASLVPLVLAARYPYGVYRDSRSTLVIHNLAHQGVPEASTYPDILGLPPEWYG 338  
Db 121 MFVVDWHASLVPLVLAARYPYGVYRDSRSTLVIHNLAHQGVPEASTYPDILGLPPEWYG 180  
QY 339 ALEWVPEWARRHALDKGEAVNFKAGVAVTADRIYTVSGYSGWVTTABGGOGNELLS 398  
Db 181 ALEWVPEWARRHALDKGEAVNFKAGVAVTADRIYTVSGYSGWVTTABGGOGNELLS 240  
QY 399 RKSVLNGIYVNGIDINDWNPATDKCIPCHYSVDLGSKAKCAELKELGLPVRDVLIG 458  
Db 241 RKSVLNGIYVNGIDINDWNPATDKCIPCHYSVDLGSKAKCAELKELGLPVRDVLIG 300  
QY 459 FGRLDYQKIDILKMAIPELMREDVQFVNLGSGDPIFEGWMRSTESSYKDKFGRWVGF 518  
Db 301 FGRLDYQKIDILKMAIPELMREDVQFVNLGSGDPIFEGWMRSTESSYKDKFGRWVGF 360  
QY 519 VVSHRITAGCDILLMPSEFPCGLNOLYAMQYGVVHVHGTGGLRDVTFENFNFGAKE 578  
Db 361 VVSHRITAGCDILLMPSEFPCGLNOLYAMQYGVVHVHGTGGLRDVTFENFNFGAKE 420  
QY 579 EGTGNAFSLPTVDMK 593  
Db 421 EGTGNAFSLPTVDMK 435

RESULT 5  
US-09-606-304-10  
Sequence 10, Application US/09606304  
GENERAL INFORMATION:  
APPLICANT: Kossmann, Jens  
Springer, Franziska  
Abel, Gernot

TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESS: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS



```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/606,304
; FILING DATE: 28-Jun-2000
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,567
; FILING DATE: <Unknown>
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
;
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
;
; PS-09-606-304-10

```

Query Match	49.38;	Score 1992;	DB 20;	Length 641;
Best Local Similarity	67.59;	Pred. No. 3.9e-164;		
Matches 360;	Conservative 80;	Mismatches 87;	Indels 6;	Gaps 4;
QY 102	LLLEGIADS--IDSIIVAASEQD--SEIMDANEQPOAKVTRTSIVFTGEEAAPYAKSGGL	157		
DB 90	LPHSVAGDAGTWESHDIVANDROOLSDSDTEEMETPIKLTNFIIVTAAAPYSKTGGL	149		
QY 158	GDVCSGLPIALAARHRYVMYVMPYRLNGS--SDKNYAKALYAKTHIKIPCGSGHSEVTFPH	216		
DB 150	GDVCSGLPMALAARHRYVMYVSPRYLNGSPSDEKYANAVDLDRATRVHCFGDAQAEVAFYH	209		
QY 217	EYRDNVDMVFDHPSYHRPGSLYGDNGFCAGDGNQFRYTLILCYAACCAEAPLILELGGYIYQG	276		
DB 210	EYRAGVDMVFDHSSYCRPGTPYGDIIYCAEFGDGNQFRFTLLSHAACCAEAPLVLPIGGTYGE	269		
QY 277	NCMFVYNDHHAISLVPVLLAAKYPGVYVYRDSRSLVTHNLHAGHGVBPASTYPDLGLPPPW	336		
DB 270	KCLFLANDHWAALVPLLAAKYRPGVYVKDARSVALHNTAHQVEPAPVYNNLGLPPPW	329		
QY 337	YGAEWVFPPEWARHRAKOGCAVNFPGKAVVTADRIVTVSQGYSWEVTTAEGSQGLNELL	396		
DB 330	YGAEWIFPPTWARALDGTGETVNVKGAIAVADRILTVSQGYSWEITTPEGGYGLHELL	389		
QY 397	SSRSKSYLNGIVANGIDINOWNPITDKCLPHHYSVDLGSKACKCAEKLQELGLPVRDVPVL	456		
DB 390	SSRSQSVLNGITNGIDVNDWNPSTDEHIAHSYINDLSKQVCKTDLQELGLDIRPDCLP	449		
QY 457	IGFIRGLDYQKIDILKNAIPELMPREDYQVYMLGSGDPIPEGWNRSTESSYKDKKRGWVG	516		
DB 450	IGFIRGLDYQKGVDIILSAIPELMDQNVQVYMLGSGEKQYEDMMRHTENFLFKDKFRWVG	509		
QY 517	FSPVPVSHRTAGCDIILLMPSPCEPGLNOLVAMOYGTVPVYVHGTGGGLDRDVTETFNFPFGAK	576		
DB 510	FNPVPVSHRTAGCDIILLMPSPREPCGLNOLVAMYGTIPIVHSTGGGLRDIVKDFNPYAOE	569		
QY 577	G-BEGTGAWAFSPITVDKMLWALRTAMSTFRHHKPSWEGLMKRGMTKOHTWDHA	628		
DB 570	GICGCGWTFSPSTFKIADTKLAIATSYTTHKHSWEGLMRGRGMDYSWENA	622		

RESULT 6  
US-09-606-304-4  
; Sequence 4, Application US/09606304  
; GENERAL INFORMATION:  
; APPLICANT: Kossmann, Jens

Springer, Franziska  
Abel, Gernot  
TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES  
INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC  
PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/606,304  
FILING DATE: 28-Jun-2000  
CLASSIFICATION: <Unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/836,567  
FILING DATE: <Unknown>  
APPLICATION NUMBER: DE P 44 41 408.0  
FILING DATE: 10-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: Agrevo-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000  
TELEFAX: 212-596-9090  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 459 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-606-304-4

五



	Query Match	31.58;	Score 1276;	DB 24;	Length 535;
	Best Local Similarity	50.5%;	Pred. No. 9.5e-102;		
	Matches 271;	Conservative 69;	Mismatches 161;	Indels 36;	Gaps 13;
QY	103	LLEGIAEDSDSIITVAASEQDS-EIMDANEQPOAKV--TRSIYFTVGTGAAPYAKSGGLGDV	160		
Db	9	ILEALSSSTKEV---ANEGDNVESKGENPPLAGANVMNVLVAACAPKVTGTTGLGDV	65		
QY	161	CGSLPIALAAARGHRVVMVMPRYLNGSSDKNKAKA--LYTAKHIKIPCGGSHEVTFPFREY	218		
Db	66	AGSLPKALARGHRVVMVVPVRY-----SHYADAQDIGNVKRYKVD--GQDMETVYFHSY	117		
QY	219	RDNVWVVFVDHPSY--HPRGSLYGNFGAAGDNQRFYLLCYAAACEAPLILLEGYIYQG--	276		
Db	118	IDGVDFVFIDSPNPRHLDNIYGN--REDILKRMVLFCKAAAEVPMHWPCGVCVCYGDG	174		
QY	277	NCFMVVDWHAISLPVLLAAKRYPYGVYVRDSRSTLVITHNLAAHQGVESPASTYDPLGLGPPEW	336		



Db 175 NLAFIANDHWTALLPVYLKAYYRDHGLMKYTRSVLVIHNTAHOGRPIDDFRYTDLPEHY 234  
QY 337 YGALEWVPEWARRHALDKGEAVNFKLGAVVTADRIVTVSQGSWEVTTAEGQGLNELL 396  
Db 235 IDLFKLYDPV-----GGEHFNIESAGLKAADRIVTVSHGVAMEIKTSEGCGWGLHGII 286  
QY 397 SSRKSVLNGVINGIDINDWNPTTKCLPH-----HYSVDDL-SGKAKCKAELQKELGLPVR 451  
Db 287 NENDMWLGRVINGIDITKDNPKIDVHLKSDGYNTYLETLSQSKRQCKAALQKELGLPVR 346  
QY 452 EDVPLGFGTGRLOYQKIDILKMAIPELMREDVQFVNLGSGDPIFGGWRMSTESSYKDKF 511  
Db 347 EDVPLGFGTGRLOYQKIDILKMAIPELMREDVQFVNLGSGDPIFGGWRMSTESSYKDKF 511  
QY 512 RGWVGSVPVSHRITAGCDILMPSRFEPCGLNQLYAMQYGTVPVHGTGGLRDTVETFN 571  
Db 407 RGWVGSVPVSHRITAGCDILMPSRFEPCGLNQLYAMQYGTVPVHGTGGLRDTVETFN 571  
QY 572 PFCAKEEGTGWAFSPVLDVDMKMLWALRTAMSTFREHKPSWEGMLKRGMTKDHWDHA 628  
Db 467 PF---BESGLGWTFSATNKLINALGNCLLTFRQYKQSGWGLRQRGMTOQLSDWNA 520  
RESULT 10  
US-09-402-254-51  
; Sequence 51, Application US/09402254  
; GENERAL INFORMATION:  
; APPLICANT: Keeling, Ganping  
; TITLE OF INVENTION: PLANT LIKE STARCHES AND THE METHOD OF MAKING THEM IN  
; TITLE OF INVENTION: HOSTS  
; FILE REFERENCE: 2461-52  
; CURRENT APPLICATION NUMBER: US/09/402,254  
; CURRENT FILING DATE: 1999-10-01  
; EARLIER APPLICATION NUMBER: PCT/US98/06660  
; EARLIER FILING DATE: 1998-04-03  
; EARLIER APPLICATION NUMBER: 60/042,939  
; EARLIER FILING DATE: 1997-04-04  
; NUMBER OF SEQ ID NOS: 77  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 51  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-402-254-51  
Query Match 30.3%; Score 1225.5; DB 18; Length 670;  
Best Local Similarity 41.2%; Pred. No. 3.5e-97; Mismatches 97; Indels 81; Gaps 16;  
Matches 277; Conservative 97; Mismatches 97; Indels 81; Gaps 16;  
QY 12 APSVRLRADPATARASACVVRARLRAR-LGRY---VAELSREGPAARPAQQQLAPPL 67  
Db 10 APPER-SGDAARLPARRNNAVSRDPLQPVGRYSGATGTATGAAS--CQNAADVE 66  
QY 68 VPQFLAPPPAPAPAP-TQPLPDAGVGEALAPDLLL-----EGIAED 110  
Db 67 IKSIVAAPTSTIVKFPAPGYRMLPS---GDIAPETVLPAPKPLHESPAVDGSDNGIAPP 123  
QY 111 SIDSIIIVASEQDSEIMDANEQOAK-----VTRSIIVF 143  
Db 124 TVEPLVQEAATWDPKYYIGDEPEDEKDDSRVGDAGSFEHYGNDSPGAGENVMNIV 183  
QY 144 VTGEAAPAYAKSGGLGVDCSLPTALAAHRVNMVMPYRNLGSSDKNYAKALYAKHIKI 203  
Db 184 VRAECSFWCKTGGGLGVGALPALARRHVRNMVVPY-----GDYVEAFDMGIRKYY 237  
QY 204 PCGGSHVETTFHEYRDNDVWVVDHPSY-HRPGSLYGNFGAFGDNQFRYTLCLYAAACE 262  
Db 238 KAAGQDLVNYHFAFTDGVDFVFDAPLFRHRODDIYG---GSRQETMKRMILFCRAVE 294  
QY 263 APLILELGVYIGO-NCMFVNDWHDASLVPLLAARYPGVYRDSRSTLVIHNLAHQGV 321

Db 295 VPHVPCGVCYGDGNLVFTIANDHWTALLPVYLKAYYRDHGLMAYTRSVLVIHNTAHQGR 354  
QY 322 EPASTYPTDLGPPEWYGALEWVPEWARRHALDKGEAVNFKLGAVVTADRIVTVSQGSY 381  
Db 355 GPVDERPYMDPEHYLQHFELYDPV-----GGEHANIFAAGLKMADRVVTVSRGYLW 406  
QY 382 EVTTAEGGQGLNELLSSRKSVLNGVINGIDINDWNPTTKCLPH-----HYSVDDL-SGKA 436  
Db 407 ELKTVGEGWGLHDIIRSNWKNKINGIVNGIDHQNPKVDVHLKSDGYNTYLETLDAGKR 466  
QY 437 KCAELQKELGLPVRREDVPLIGFTGRLOYQKIDILKMAIPELMREDVQFVNLGSGDPIF 496  
Db 467 QCAALQRELEGLVREDVPLIGFTGRLOYQKIDILKMAIPELMREDVQFVNLGSGDPIF 496  
QY 497 EGMWSTESSYKDKFRGWGVSFVPSHRTITAGCDILMPSRFEPCGLNQLYAMQYGTVPV 556  
Db 527 ERMQLHLERHPNKRGVGWGVSFVPSHRTITAGCDILMPSRFEPCGLNQLYAMQYGTVPV 586  
QY 557 VHTGGLRDTVETFNPFCAKEEGTGWAFSPVLDVDMKMLWALRTAMSTFREHKPSWEGMLK 616  
Db 587 VHAAGGLRDTVAPDFP---GDAGLGWTFDRAEANKLIEALRHCLDITRYKYGESWKSQA 643  
QY 617 RGMTKDHWDHA 628  
Db 644 RGMSQLSDWHA 655  
RESULT 11  
US-09-606-304-8  
; Sequence 8, Application US/09606304  
; GENERAL INFORMATION:  
; APPLICANT: Kossmann, Jens  
; APPLICANT: Springer, Franziska  
; APPLICANT: Abel, Gernot  
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES  
; INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC  
; PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/606,304  
; FILING DATE: 28-Jun-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/836,567  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: DE P 44 41 408.0  
; FILING DATE: 10-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley Jr., James F.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: Agrevo-4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-596-9000  
; TELEFAX: 212-596-9090  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 767 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:



US-09-606-304-8

Query Match  
Best Local Similarity 30.0%; Score 1215; DB 20; Length 767;  
Matches 264; Conservative 78; Mismatches 188; Indels 48; Gaps 12;  
QY 79 PAQSPAPTOPPL-----PDAGVGEAPDLLEGIA-----EDSIDSIIAAS 120  
DB 195 PPEPKSQTLLDVNSRKSLSVDPGKIQSYMPSLKRKSSASHVEORNELENSSAEAN 254  
QY 121 EODSEIMDANPOQAKV---TRSTIVFTGEAAPYAKSGGLDVCGLSPIALAARGHRVMV 177  
DB 255 EETEDPVNIDEKPPPLAGTNMNIILVASECAPWSKTGGLDVGALPKALARRGHRVMV 314  
QY 178 VMPRYLSSGDKNYAKALYAKHILKIPCFGSGSHEVTFEYHRDNDVDFVD-HPSYHRPG 236  
DB 315 VAPRYDNPQDSG-----VRKIYKVD--GODVEVTFQAFIDGVDFVFDISHMFRHIGN 368  
QY 237 SLXGDNFAGDNQFRYTLACYACEAPLILELGGYIYGQ-NCMFVNDWHSALVPVLLA 295  
DB 369 NIYGN---RVDILKRMVLFCKAAIEVPHVPCGGVCYCGNLFVFIANDHHTALLPYLK 425  
QY 296 AKRYPGYVDRSRLVIHNLHOGVEPASTYPDLGLPPEWYGALEWVFFEWARRHALDK 355  
DB 426 AYYRDNIMYTRSVLVIHNLHOGGRPLEDFSYVDLPPHYMDPFKLYDPV-----G 477  
QY 356 GEAVNFKGAVTADRVTVSQYSWEVTTAEGGOLNELLSSRKSVLNGIVNGIDINDW 415  
DB 478 GEHFNFAAGLKTADRVTVSHGYSWELKTSQGGWGLHQIINENDWKLQGIIVNGIDTKEW 537  
QY 416 NPTDKCLPH-----HYSVDDL-SGKAKCAELQKELGVPVREDVPLIGFIRLDYQKID 470  
DB 538 NPELDVLHQSOGYMNYSILDTLOTQCKPOCKAALQKELGLPVRDQVPLIGFIRLDYQKID 597  
QY 471 LIKMAIPELAREDOVFMVLCGSDPILEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCD 530  
DB 598 LIAEASAMWQDVQLVNLGTGRDLQMLRQECQHNDKIRGWGFSVKTSHRITAGAD 657  
QY 531 ILLMPSRFEPCGLNQLYAMQYTPVHVHGTGGLRDTVETENPFGAKGEETGWAFSPLTV 590  
DB 658 ILLMPSRFEPCGLNQLYAMQYTPVHVHGTGGLRDTVETENPFGAKGEETGWAFSPLTV 590  
QY 591 DKMLWALRTAMSTFREHKPSWGLMKRGMTKDHWDHA 628  
DB 715 SOLIHALGNCLTYREYKKSWEIGIOTRCMTQDLSWDNA 752

RESULT 12  
US-09-402-254-49  
; Sequence 49, Application US/09402254  
; GENERAL INFORMATION:  
; APPLICANT: Guan, Hanning  
; TITLE OF INVENTION: PLANT LIKE STARCHES AND THE METHOD OF MAKING THEM IN  
; FILE REFERENCE: 2461-52  
; CURRENT APPLICATION NUMBER: US/09/402,254  
; EARLIER FILING DATE: 1999-10-01  
; EARLIER FILING DATE: 1998-04-03  
; EARLIER FILING DATE: 1997-04-04  
; NUMBER OF SEQ ID NOS: 77  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 49  
; LENGTH: 804  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-402-254-49  
Query Match 29.9%; Score 1209.5; DB 18; Length 804;

Best Local Similarity 41.9%; Pred. No. 1.2e-95;  
Matches 288; Conservative 85; Mismatches 228; Indels 87; Gaps 18;  
QY 12 APSVRLRADPA-----TAARASACVVRRLRLARGRYVAELSPREGPAAPAOQOQLAPP 66  
DB 56 AASVRAAAAAGSESEAAKSSSS-SQAGAVQGSTAKAVDSASPPNPLTSAPKOSASAM 114  
QY 67 L--VPGEFLAPPPAPAPQSPAPTQPLP-----DAG-----VGLAPDLILLEGIABDSI 112  
DB 115 QNGTSGSSASTAAPVSGPKADHPSAPVTKREIDASAVKPEPAGDDARPVESIGIAE-PV 173  
QY 113 DSIIVASEODS-----EIMDANEQPOA-KVTRSVIVFTGEAAPYAKSGGLDVCGLSP 165  
DB 174 DAKADRAPATDAASAPYREDNEPGLAGPNVNVVVVAVASECAPCKTKTGLDGVVVALP 233  
QY 166 IALAARGHRVMVMPRYLNGSSDKNYAKALYAKHILKIPCFGSGSHEVTFEYHRDNDVW 225  
DB 234 KALARRGHRVMVPIRY-----GEYAEARDLGVRRRYKVAGQDSEVTVYFHSIDGVDFV 287  
QY 226 FVDHPSY-HRPGSLYGLDNFGAGDNQFRYTLACYACEAPLILELGGYIYGQ-NCMFVNV 283  
DB 288 FVEAPPFRHNNIYG---GERLDILKRMILFCKAAVEVPYAPCGGTVYVYDGNLVFIAN 344  
QY 284 DHASIVPVLAAKRYRPGYVDRSRLVIHNLHOGVEPASTYPDLGLPPEWYGALEWV 343  
DB 345 DHTALLPYLKYAYRDNGLMQYARSVLVIHNLHOGRPVDFVDFVDFVDFVDFVDFVDFV 394  
QY 344 FPEWARRHALDK-----GEAVNFKGAVTADRVTVSQYSWEVTTAEGGOLNELL 395  
DB 395 -----EHYIDHFKLYDNIGDHSNFAAGLKTADRVTVSNYMGWELKTSSEGGWGLHDI 448  
QY 396 LSSRKSVLNGIVNGIDINDWNTDTCCLPH-----HYSVDDL-SGKAKCAELQKELGLPV 450  
DB 449 INDNWKLQGIIVNGIDIMSEWNPVAVDVLHSHDDYNTFTFDTGKROCKAALQRLGLQV 508  
QY 451 REDVPLIGFIRLDYQKIDLIKMAIPELAREDOVFMVLCGSDPILEGWMRSTESSYKDK 510  
DB 509 RDDVPLIGFIRLDYQKIDLIKMAIPELAREDOVFMVLCGSDPILEGWMRSTESSYKDK 568  
QY 511 FRCWGFSPVSHRITAGCDILLMPSRFEPCGLNQLYAMQYTPVHVHGTGGLRDTVET 570  
DB 569 VRAWGFSPVSHRITAGCDILLMPSRFEPCGLNQLYAMQYTPVHVHGTGGLRDTVET 628  
QY 571 NPFGAKGEETGWAFSPLTVDKMLWALRTAMSTFREHKPSWGLMKRGMTKDHWDHAP- 629  
DB 629 DPF---NDTGLGWTDFRAEANRMDALSCLTYRYNKESWRACRARGMAEDLSWDHAAV 685  
QY 630 -----SSTSRSSSGSPSWT 642  
DB 686 LYEDVLVKAKYQWANLATRRRCRTWT 713

RESULT 13  
US-09-625-406-11  
; Sequence 11, Application US/09625406  
; GENERAL INFORMATION:  
; APPLICANT: Keeling, Peter  
; APPLICANT: Guan, Hanning  
; TITLE OF INVENTION: Starch Encapsulation  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:



; APPLICATION NUMBER: US/09/625,406  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA: 08/941,445  
; APPLICATION NUMBER: 08/941,445  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winner, Ellen P  
; REGISTRATION NUMBER: 28,547  
; REFERENCE/DOCKET NUMBER: 89-97  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 698 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-625-406-11

Query Match 29.8%; Score 1205.5; DB 20; Length 698;  
Best Local Similarity 43.0%; Pred. No. 2.1e-95;  
Matches 283; Conservative 83; Mismatches 221; Indels 71; Gaps 17;  
  
QY 12 APSVRLRADPA-----TAAASACVVRARLRLARGRYVAELSREGPAARPAQOQOLAPP 66  
Db 56 AASVRAAAAPAGSEAEAKSSSS--SQAGAVQSGTAKAVDSAPPNPLTAPKQSQSAA 114  
QY 67 L--VPGFLAPPPAPASPAPTQPLP-----DAG-----VGLAPDLLLEGIADSI 112  
Db 115 QNGTSGSSASTAAPVSGPADHPSAPVTKREIDASAVKPEPAGDARPVESIGIAE--PV 173  
QY 113 DSIIVASEQDS-----ETMDANEQQA-KVTRSIIVFTGEAAPYAKSGGLGDCVCSLP 165  
Db 174 DAKADAAPATAAASAPYREDNEPGLAPGNVNMVVVVAASECAPFCKTGGLGVDVGCALP 233  
QY 166 IALAARGHRVMVMPRYLNGSSDKNYAKALYAKHIKIPCGGSHSEVTFEHEYRDNDVW 225  
Db 234 KALARGHRVMVIPRY-----GEYAEARDLGVRRYKVKAGQDSEVTYFHSYIDGVDF 287  
QY 226 FVDHPSY--HRPGSLYGNFAGFQNGRYTLLCYAAEAPLLELGGYIYGO--NCMEVFN 283  
Db 288 FVEAPPFRHRHNTYG---GERLDLKRMLFCCKAAVEVPMYAPCGGTIVYGDGNLFIAN 344  
QY 284 DWHASLPVLLAAKRYPGVYRDSRSTLVINLHAGQVEPASTYPDGLGPPWYGALEWV 343  
Db 345 DWHALLPVLYKARYRDNLQYARSVLVINIAHQGRGVPDDFVDFDLP-----394  
QY 344 FPEWARRHALDK-----GEAVNFKGAVVTADRVITYSQGSWEVTTAEGGQGLNEL 395  
Db 395 -----EHYIDHFKLYDNIGDHSNVFAAGLKTADRVVTVSNGYMWELKTSEGGWGLHDI 448  
QY 396 LSSRKSVLNGIVNGIDINDWNTTDKCLPH----HYSVDDL--SKAKCKAELOKELGLPV 450  
Db 449 INQNDWKLOGIVNGIDSEWNPADVHLHSDDYNTYFETDITGRKQCKAALQOLGLQV 508  
QY 451 REDVPLIGFGRLDYQKIDILKMAIPELMREDVQFVLMGSDPIFEGWNRSTESSYKDK 510  
Db 509 RDDVPLIGFGRLDHQKGVDIADAIHWIAGQDVQLVLMGTGRADLEDMLRFESEHSK 568  
QY 511 PRGWVGFSPVSHRITAGCDILLMPSPFPCGLNQLYAMQYGTVPVVGHTGGLRDTVET 570  
Db 569 VRWVYGFSPVLAHRTAGADILLMPSPFPCGLNQLYAMAYGTVPVVAVGGLRDTVAPF 628  
QY 571 NPEGAKGEGGTGWAFFSLPTVSKMLWALRTAMSTFREHKPSWEGMLKRGMTKDHTWDA 628  
Db 629 DPF-----NDTGLGWTFDRAEHRMIDALSHCLITTYRNYKESNRACRARGMAEDLSWDHA 683

RESULT 14  
US-09-388-743-26

; Sequence 26, Application US/09388743  
; GENERAL INFORMATION:  
; APPLICANT: Singletary, George  
; APPLICANT: Zhou, Lan  
; TITLE OF INVENTION: Novel Starch Synthase Polynucleotides and Their  
; TITLE OF INVENTION: Use in the Production of New Starches  
; FILE REFERENCE: 1144  
; CURRENT APPLICATION NUMBER: US/09/388,743  
; CURRENT FILING DATE: 1999-09-02  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 26  
; LENGTH: 801  
; TYPE: PRT  
; ORGANISM: Typha latifolia  
; US-09-388-743-26

Query Match 29.6%; Score 1196.5; DB 17; Length 801;  
Best Local Similarity 42.6%; Pred. No. 1.6e-94;  
Matches 269; Conservative 80; Mismatches 200; Indels 83; Gaps 15;  
  
QY 42 GRVVAELSRGPAARPAQOQOLAP-----PLVPGFLAPPPP-----APAQSAPT 86  
Db 109 GKYS-----GAVPNYSQLAODDTSENPLVNSFGGSPKDNVEAVEFQVRQSAVDAG 161  
QY 87 QPPLPDAGVAGELAPDLLLEG-----TAEDSIDSIIVAASQDSEIMDANEQP- 133  
Db 162 RPEPSLGTTKILSPFYLEAESDCAENAEADLVEAKLDSHVH---KDDLNPGEENEVPL 218  
QY 134 --QAKVTRSTIVFTGEAAPYAKSGGLGDCVCSLPIALAARGHRVMVMPRYLNGSSDKNY 191  
Db 219 PLAGANYMNIIVAAECAPMSKTKGGLGDVAGALPKALARHGRVMVVPARYGNTAEQDI 278  
QY 192 AKALYAKHIKIPCGGSHSEVTFEHEYRDNDVWVFDHPSY--HRPGSLYGNDFCAFQDNQ 250  
Db 279 GVRKYKVVH-----GQDMEVTFYHAYIDGVDFVMDSPDFRHRGNRIYEGN---RVDIL 329  
QY 251 FRYTLLCYAAEAPLLELGGYIYGO--NCMFVNDWHASLPVLLAAKRYPGVYRDSRS 309  
Db 330 KRMILFCKAAVEVPMVHPVPCGFCYGDGNLAFITNDNHTALLPVLYKARYDRNGLMKYARS 389  
QY 310 TLVTHNLHAGQVEPASTYPDGLGPPWYGALEWVPEWARRHALD-----KGEAVNF 361  
Db 390 VLVTHNLHAGQVEPDDFKFVGLPD-----HYLDLFLYDVPVCGEHLNI 433  
QY 362 LKGAVVTADRVITYSQGSWEVTTAEGGQGLNELLSRKSVLNGIVNGIDINDWNTTDK 421  
Db 434 FAAGLKTADRVTVSHGYAWELKTSEGGWGLHEINESNNKFGQIVNGIDAKEWSPEDV 493  
QY 422 CLPH----HYSVDDL--SKAKCKAELOKELGLPVREDVPLIGFGRLDYQKIDILKMAI 476  
Db 494 HLKSDGYTNSLDPLTEMGKPVCKAALQVGLPVRDNPVPIAFIGRLDHQKGVDLIAEAM 553  
QY 477 PELMREDVQFVLMGSDPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDILLMP 536  
Db 554 PWIVSHDVQVVMGLTGQDLLENLLRNFGQHRDKVRWVAEFSVKMAHRTAGADILLMP 613  
QY 537 RFEPCLNQLYAMQYGTVPVVGHTGGLRDTVETPNFCAKGEETGNWAFSLPTVDMKLWA 596  
Db 614 RFEPCLNQLYAMMYGTIPVVAHVAVGGLRDTVTQDFPE---NESGLGWTDFDRAEAGKLIHA 670  
QY 597 LRTAMSTFREHKPSWEGMLKRGMTKDHTWDA 628  
Db 671 LNNCLNTYWNYSKSNKGLQTRGMMQDLSWDNA 702

RESULT 15  
US-09-606-304-6  
; Sequence 6, Application US/09606304  
; GENERAL INFORMATION:  
; APPLICANT: Kossmann, Jens  
; Springer, Franziska







us-09-674-824-2.rapm

Fri Mar 29 09:26:42 2002



Sequence	Srtd Orig	ZScore	EScore	Len	Documentation
gb_pat:AX010492	+ 4044.00	3630.28	6.3e-194	2771	! AJ010492 Sequence 1 from Patient
gb_pl:TAE292522	+ 3998.00	3589.83	1.1e-191	2421	! XJ292522 Triticum aestivum mRNA
gb_pl:TAE292521	+ 3933.00	3531.11	2.1e-188	2575	! AJ292521 Triticum aestivum mRNA
gb_pl:AF091803	+ 3930.00	3528.38	3.0e-188	2591	! AF091803 Triticum aestivum stat
gb_pat:AX031272	+ 3930.00	3528.19	3.0e-188	2662	! AX031272 Sequence 11 from Patient
gb_pat:AG33350	+ 3612.00	3243.97	2.1e-172	2239	! AG33350 Sequence 11 from Patient
gb_pl:TAU48227	+ 3313.00	2976.20	1.7e-157	2055	! U48227 Triticum aestivum solub
gb_pl:AF165890	+ 2829.50	2540.63	3.1e-133	2585	! AF165890 Oryza sativa subsp. J
gb_pat:E06904	+ 2821.50	2533.60	7.6e-133	2533	! E06904 cDNA encoding soluble s
gb_pl:RICSS5	+ 2821.50	2533.60	7.6e-133	2533	! D16202 Rice mRNA for soluble s
gb_pat:AG33359	+ 2809.50	2533.25	2.9e-132	2383	! AG33359 Sequence 1 from Patient
gb_pl:AF036891	+ 2776.50	2492.05	1.6e-130	2991	! AF036891 Zea mays starch synth
gb_pl:AF168786	+ 2770.50	2487.86	2.4e-137	2592	! AF168786 Sorghum bicolor solub
gb_pat:AR106494	+ 2713.50	2439.33	1.4e-127	1752	! AR106494 Sequence 12 from Pat
gb_pat:AG3308	+ 2655.00	2383.44	1.8e-134	2992	! AG3308 Sequence 1 from Patient
gb_pat:AR049919	+ 2651.00	2379.41	2.9e-134	2990	! AR049919 Sequence 1 from Patient
gb_pl:AF234163	+ 2622.50	2346.64	2.0e-122	8387	! AF234163 Hordeum vulgare starch
gb_pat:AR106498	+ 2546.50	2289.88	2.9e-119	1620	! AR106498 Sequence 20 from Pat
gb_pat:AX031274	+ 2349.00	2059.70	1.1e-108	10336	! AX031274 Sequence 13 from Pat
gb_pl:AF091802	+ 2315.00	2068.51	6.1e-107	11387	! AF091802 Agelopsis tauschii st
gb_pat:AS1207	+ 2001.00	1797.65	7.5e-92	2360	! AS1207 Sequence 9 from Patient
gb_pat:AR112888	+ 2001.00	1797.65	7.5e-92	2360	! AR112888 Sequence 9 from Pat
gb_pl:STSTASNT	+ 2001.00	1797.65	7.5e-92	2360	! Y10416 S. tuberosum mRNA for sc
gb_pat:E0626295	- 1786.00	1574.85	1.9e-79	170371	! AB026295 Oryza sativa genom
gb_pl:RIC1E1	+ 1756.50	1568.94	4.1e-79	8900	! D38221 Rice gene for soluble s
gb_pat:AS1201	+ 1731.50	1557.81	1.7e-78	1758	! AS1201 Sequence 3 from Patient
gb_pat:AR112885	+ 1731.50	1557.81	1.7e-78	1758	! AR112885 Sequence 3 from Pat
gb_pl:AF121673	+ 1526.00	1366.63	7.6e-68	4617	! AF121673 Arabidopsis thaliana
gb_pl:AB006297	+ 1517.50	1338.77	2.7e-66	84194	! AB006297 Arabidopsis thaliana
gb_pl:AF068834	+ 1253.50	1126.38	1.8e-54	2480	! AF068834 Zea mays starch synth
gb_pat:AR106493	+ 1250.50	1120.81	3.4e-54	2325	! AR106493 Sequence 10 from pat
gb_pl:AF019296	+ 1246.00	1124.13	2.7e-53	2097	! AF019296 Zea mays starch synth
gb_pl:AF039537	+ 1239.50	1112.81	1.0e-53	2865	! AF039537 Oryza sativa solub
gb_pat:AS1205	+ 1215.00	1090.99	1.7e-52	2793	! AS1205 Sequence 7 from Patient
gb_pat:AR112887	+ 1215.00	1090.99	1.7e-52	2793	! AR112887 Sequence 7 from Pat
gb_pl:AF026421	+ 1212.50	1088.14	2.5e-52	3049	! AF026421 Chlamydomonas reinhar
gb_pat:TAE269503	+ 1204.50	1081.60	5.7e-52	2780	! TAE269503 Triticum aestivum mRN
gb_pl:PSSTASNT	+ 1204.50	1080.65	6.5e-52	2998	! X88790 P. sativum mRNA for star
gb_pl:TAE269504	+ 1203.50	1080.67	6.4e-52	2793	! TAE269504 Triticum aestivum mRN



380 CGCGGCTCCGCGCTTGGCGGGGGCCGCTACGTCGCGGAGCTCAGCAGG 429  
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67  
430 GAGGCT 479  
67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84  
480 CGTGGCAGGCTTCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 529  
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100  
530 CCCCAGCAGCGCCCGCTCGCGGACGCGCGCGCGCGCGCGCGCGCG 579  
101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117  
580 GACCTCCCTCGTGAAGGATTTGCTGAGGATTTCCATCGACAGCATAAATGT 629  
117 IAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134  
630 GGCTGCAAGTGAGCAGGATTTCTGAGATCATGTGATGCGGATGCGCAACCTC 679  
134 IAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150  
680 AAGCTAAAGTTACAGTACGATCGCTGTTGCTGACTGCTGAAGCTGCTCCT 729  
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167  
730 TATGCAAAAGTCAAGGGGCTGGAGATGTTGTTGCTGCTGATACCAATTTGC 779  
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184  
780 TCTTGTCTGCTGCTGCTGCTACCGAGTATGTTGTTGATGCCAAGATACTTAA 829  
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200  
830 ATGGGTCTCTCATAAAACTATGCAAGGCATTTATACACTGCGAAGCAC 879  
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG1 217  
880 ATTAAGATTTCCATGCTTTTGGGGATCATCATGAAGTGACCTTTTTCATGA 929  
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234  
930 GTATAGAGACAACGTCGATGGTGTGTTGTCGATCATCGTCATATCACA 979  
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250  
980 GACCAGGAAGTTTATATGGAGATAAATTTGGTGTCTTTGGTGATATATCAG 1029  
251 PheArgTyrThrLeuLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267  
1030 TTCAGATACACACTCTTTGCTATGCTGATGCGGAGCGCCCGCCACTAATCCT 1079  
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284  
1080 TGAATTTGGGAGGATATATTTATGACAGAAATTCATGTTTGTGTGAAGC 1129  
284 sprPheHisAlaSerLeuValProValLeuLeuAlaLysTyrArgPro 300  
1130 ATTGGCATGCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1179  
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317  
1180 TACGGTGTTTACAGAGATTTCCCGCAGCACCCCTTGTATATACATAAATTTAGC 1229  
317 aHisGlnGlyValIleProAlaSerThrTyrProAspLeuGlyLeuProP 334  
1230 ACATCAGGGGTGTGAGCGCTCAAGTACATATCCTGATCTGGGATTCGCTC 1279  
334 roGluTyrTyrGlyAlaLeuGluTyrValPheProGluTyrPheAlaArgArg 350  
1280 CTGAATGCTATGGAGCTTTAGATGGGTATTTCCAGAAATGGGCAAGGAGG 1329

351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367  
1330 CATGCCCTTTGACAAAGGCTCAGGAGTTAACTTTTGAAGGAGCAGTTGT 1379  
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384  
1380 GACAGCAGATCGGATTTGACCGTCTCAGTCAGGGTTATTTCATGGGAGTCA 1429  
384 hrThrAlaGluGlyGlnGlyLeuAsnGlnLeuLeuSerSerArgLys 400  
1430 CAACCTGCTGAAGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCCGAAAA 1479  
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417  
1480 AGTGTATTGAATGCAATTTGTAATGCAATTTGACATTTAATGATTGGAACCC 1529  
417 oThrThrAspLysCysLeuProHisIleThrSerValAspAspLeuSerG 434  
1530 CACCACAGACAGTGTCTCCCTCATCATTTATCTGTCGATGACCTCTCTG 1579  
434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450  
1580 GAAAGGCCAAATGTAAGCTGAATTCAGAGGAGGTTGGGTTTACCTGTA 1629  
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467  
1630 AGGAGGATGCTCTCTGATTTGGCTTTTATTTGGAAGACTGSAATTTACCAGAA 1679  
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484  
1680 AGGATTGATCTCATTAATAATGCCCATTCAGAGCTCATGAGGGAGGACG 1729  
484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTyrMet 500  
1730 TGCAATTTGTCATGCTTGGATCTGGGATCCAAATTTTGAAGGCTGGATG 1779  
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517  
1780 AGATCTACCGAGTCGAGTTACAAGGATAAATTCGCTGGATGGGTGGATT 1829  
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534  
1830 TAGTGTCCAGTTTCCACAGATAAATTCGAGGTTGCGATATATTGTTAA 1879  
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550  
1880 TGCCATCGAGATTTGAACCTTTCGCTCTTAATCAGCTATATGCTATGCAA 1929  
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567  
1930 TATGTCACAGTTCTCTGTAGTTTCATGGAACTGGGGGCTCCGAGACACAGT 1979  
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584  
1980 CGAGACCTTCAACCTTTTGTGTGCAAAAGGAGAGGAGGTACAGGGTGGG 2029  
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600  
2030 CGTTCACCGCTAACCGTGCACAGATGTTGTGGGCATTCGGAACCCCG 2079  
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617  
2080 ATGTCACACATTCAGGAGCACACAGCCGCTCTGGGAGGGGCTCATGAAGCG 2129  
617 gGlyMetThrLysAspHisThrTrpAspHisAlaProSerSerThrSera 634  
2130 AGGCATGACGAAAGACCATACGTGGGACCATCCCGGAGCAGTACGAGCA 2179  
634 rgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgGly 650  
2180 GATCTTCGAGTGGGCTTCGTGGGACCAACCCCTACGTCATGTAGAGGGGA 2229







627 TCTTGCTGCTGCTGCTCACCGAGTGATGGTTGTAATGCCAAGATACCTTAA 676  
184 sNGlySerSerAspLysAsnTyrAlaLysAlaLysAlaLysTyrThrAlaLysHis 200  
677 ATGGGCTCTCTGATTAATAAACTATCAAAAGGATATATACACTGCGAAGCAC 726  
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG1 217  
727 ATTAAGATTCCATGCTTTGGGGATCATCATGAAGTGAACCTTTTTCATGA 776  
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234  
777 GTATAGACAACTGATGGTGGTGTTCGATCATCCGCTCATATCACA 826  
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250  
827 GACCAGGAAGTTTATATGGAGATAATTTGGTGCTTTTGGTGTAATCAG 876  
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267  
877 TTCAGATACACACTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926  
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284  
927 TGAATTGGGAGGATATATTTATGGACAGATTTGATGTTTGTGTGAACG 976  
284 spTrpHisAlaSerLeuValProValLeuLeuAlaLysTyrArgPro 300  
977 ATTGCAATGCCACCTCTGTCAGGCTCTCTCTCTCTCTCTCTCTCTCTCT 1026  
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317  
1027 TAGGGTGTACAGAGATCCCGCAGCACCTTGTGTATACATAATTTAGC 1076  
317 ahisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProP 334  
1077 ACATCAGGGTGTGGAGCTGCAAGTACATATCTCTGATCTGGGATTCGCTC 1126  
334 toGluTyrTyrGlyAlaLeuGluTyrPheProGluTyrAlaArgArg 350  
1127 CTGAATGGTATGGAGCTTTAGNATGGGTATTTCCAGATGGCAAGGAGG 1176  
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367  
1177 CATGCCCTTGACAAGGGTGAGCAGTTAACTTTTGAAGGAGCAGTTGT 1226  
367 ThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValI 384  
1227 GACAGCATCGGATGTGACCGTCAGCTCAGGTTTATTCATGGGAGGTCA 1276  
384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400  
1277 CAACGTGCAAGGTGACAGGGCTCAATAGCTCTTAAGCTCCCGAATA 1326  
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417  
1327 AGTGATTAATGGAATGTAATGGAATGGAATGGAATGGAATGGAATGGA 1376  
417 oThrThrAspLysCysLeuProHisHistyrSerValAspLeuSerG 434  
1377 CACCACAGACAGTGTCTCCCTCATCATATTTCTGTCGATGACCTCTCTG 1426  
434 lYlYsAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450  
1427 GAAAGGCCAAATGTAAGCTGAATTCAGAGAGGTTGGGTTTACCTGTA 1476  
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467  
1477 AGGGAGGATGTTCTCTGATTTGCTTTTATTTGAAGACTGGATTACCGAA 1526  
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspY 484  
1527 AGGCATTTGATCATTAATAATGCCATTCAGAGCTCATGAGGAGGAGG 1576

484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500  
1577 TGCAATTTGTCATGCTTGGATCTGGGGATCCCAATTTTGAAGGCTGGATG 1626  
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517  
1627 AGATCTACCGAGTCGAGTTACAAGGATAAATCCGTTGGATGGGTTGGATT 1676  
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534  
1677 TAGTGTTCAGATTTCACACAGATAACTGCAGGTTGCGATATATGTAA 1726  
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550  
1727 TGCCATCGAGATTGAACTTGGCTTAAATCAGCTATATGCTATGCAA 1776  
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567  
1777 TATGGTACAGTTCTCTGTAGTTTCATGGAACCTGGGGGCTCCGAGACACAGT 1826  
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584  
1827 CGAGACCTTCAACCTTTTGGTGCATAAAGGAGAGGAGGTACAGGTGGG 1876  
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600  
1877 CGTTCTCACCGCTAACCGTGGACAAGATGTTGTGGCATTTGCCAACCAGCG 1926  
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617  
1927 ATGTCGACATTCAGGAGCACAGCCGCTCTGGGAGGGGCTCATGAACCG 1976  
617 qGlyMetThrLysAspHisThrTrpAspHisAla.ProSerSerThrSer 633  
1977 AGGCATGACGAAAGACCATACGTGGGACCATGCGCCGCGAGCAGTACGAGC 2026  
634 ArgSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG1 650  
2027 AGATCTTCGAGTGGGCTCTGTCGACCAACCCCTACGTCATGTAGACGGG 2076  
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667  
2077 ACTGGGAGGTCCCAAGTCCGAGTCTCTCTTGAGCTCTGAAGACATCCTCTT 2126  
667 erSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAla 683  
2127 CATCCTCCGCGGCGCGAAGGATACCCCTGTACATTTGCCGTTGTCTGCT 2176  
684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgTh 700  
2177 ACAGTAGAGTCGCAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2226  
700 rTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnL 717  
2227 ATATGACGCTGTGCTGCTGCGCGGTGACAGCTTCGGGTGGAGCAGCAGT 2276  
717 euGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAla 733  
2277 TACAGTTTGGGGGATAAGGAAGGATGCTGTCAGGATGGTTAACAGCA 2326  
734 LysHisHisSerAspGlySerLeuSerValArgValThrAlaGluIleAr 750  
2327 AAGCACCACCTCAGATGGCAGCTCTCTCTGCTGCTGCTGCTGCTGCTGCT 2376  
750 asnGlnLeuValThrLeu 756  
2377 AAACCAACTGGTGACTCTT 2395

seq\_name: gb\_pl:TAE292521

seq\_documentation\_block:

LOCUS TAE292521 2575 bp mRNA PLN 21-JUL-2000



```

DEFINITION   Triticum aestivum mRNA for starch synthase I-1 (wSSI-1 gene).
ACCESSION    AJ292521
VERSION      AJ292521.1 GI:9369333
KEYWORDS     starch synthase I-1; wSSI-1 gene.
SOURCE       bread wheat.
ORGANISM     Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Triticum.
1 (bases 1 to 2575)
Peng, M., Hucl, P. and Chibbar, R.N.
Isolation, characterization and expression analysis of starch
synthase I from wheat (Triticum aestivum L.)
Unpublished
2 (bases 1 to 2575)
Chibbar, R.N.
Direct Submission
Submitted (18-JUL-2000) Chibbar R.N., Cereal Biotechnology, Plant
Biotechnology Instt., Natl. Research Council Canada, 110 Gymnasium
Place, Saskatoon, Saskatchewan, S7N 0W9, CANADA
Location/Qualifiers
1. 2575
/organism="Triticum aestivum"
/cultivar="Fiedler"
/db_xref="taxon:4565"
/dev_stage="12-14 days-post-anthesis"
/tissue_type="developing kernels"
223..2290
/gene="wSSI-1"
223..2166
/gene="wSSI-1"
/EC_number="2.4.1.21"
/function="essential for starch synthesis"
/codon_start=1
/product="starch synthase I-1"
/protein_id="CAB9209.1"
/db_xref="GI:9369334"
/translation="MAATGVAGCLAPSRLRADPATAARASACVVRRLRLARGRY
VAELSRGPAARPAQQQLAPLVFGLAPPAPPAQSPAPTOPPLPDAGVGLAPDL
LLEGTAEDSIDIIVAASQDEIMDANKQAKVTRISIVFTGEAAYAKSGGLDV
CGSLPTAALAHGRVVMVPRYLNGSDKNYAKLYTGKHKIPGFSHVFTFHEY
RONVDWVFVDHPSYHRPGSLYGNFCAGDNQFRVTLICYACEAPLILELGGYLYGQ
NCMFVNDWHAISLVPVLLAARYPYGVYRDSRTLVIIHLAHQGVPEASTYDPLGLP
ENYGALEWFFPEARRHALDKGEAVNFKGAVVTADRIVTVSQGYSWEVTTAEGQGL
NELLSRKSVLNGIVNGIDINDNPTTKLPHYSVDDLKGAKCKAELOKELGPV
REDVPLTIGRLDYQKGIIDLKMAIPELMREDVQFVMLGSDPIFEQWMRSTESSYK
DFRQWVGFSVPVSHRTAGCDLILMPSRFPCGLNQLYAMQYGPVYVHCTGGLRDT
VETFPFGAKGEGTGWAFSPILTVDKMLWALRTANSTFREHKPSHEGLMKRGMKRDHT
WDHAAEOYEOIFEMAFVDQPYVM"
transit_peptide 223..346
/gene="wSSI-1"
/evidence=experimental
347..2163
/gene="wSSI-1"
/evidence=experimental
/product="starch synthase I-1"
2285..2290
/gene="wSSI-1"
BASE COUNT 567 a 654 c 748 g 606 t
ORIGIN

alignment_scores
  Quality: 3933.00      Length: 757
  Ratio: 5.251          Gaps: 1
  Percent Similarity: 98.943      Percent Identity: 98.547

alignment_block
US-09-674-824-2 x TAE292521
Align seg 1/1 to: TAE292521 from: 1 to: 2575
1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17

```

```

|||||
223 ATGGCGGACGAGGGCGTGGCGCGGGTGCCTCGCCCCACGCGTCCGCT 272
|||||
17 uArgAlaAspProAlaThrAlaAlaAtqAlaSerAlaCysValValArgA 34
|||||
273 GCGGCGCATCCGGAGCGGGCGCGCGCTCCGCTCGCTGCTGCGCGG 322
|||||
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
|||||
323 CGCGCTCCGCGCTTGGCGGGCGCTACGTCGCGAGTCAGCAGG 372
|||||
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProLe 67
|||||
373 GAGGCGCGCGCGCGCGCGCGCGCAGCAGCAACTGGCGCGCGCT 422
|||||
67 uValProGlyPheLeuAlaProProAlaProAlaGlnSerProA 84
|||||
423 CGTCCAGGCTTCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 472
|||||
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
|||||
473 CCGCGACGCGCGCGCTCGCGGACGCGCGTGGGGAACTCGCGGCC 522
|||||
101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleaspSerIleIleVa 117
|||||
523 GACCTCTGCTCGAAGGGATGCTGAGGATTCATCGACAGCATAAATTGT 572
|||||
117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
|||||
573 GGCTGCAAGTGAGCAGGATTCGAGATCGGATGCGAATGAGCAACCTC 622
|||||
134 lnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
|||||
623 AAGCTAAGTTACACGTAGCATCGTGTGTGACTGCTGTAAGCTGCTCCT 672
|||||
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
|||||
673 TATGCAAGTCAGGGGGCTGGGAGATGTTTGGTTCGTACCAATTGC 722
|||||
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
|||||
723 TCTTGTGCTCGTGGTCAACCGTGTGTTGTAATGCCAAGATACTTGA 772
|||||
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
|||||
773 ATGGGTCTCTGATAAAACTATGCAAGGCATATTACACTGGGAAGCAC 822
|||||
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisGl 217
|||||
823 ATTAAGATTCCATGCTTTGGGGATCACATGAAGTGACCTTTTTCATGA 872
|||||
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
|||||
873 GTATAGAGACAACGTCGATTGGGTGTTTTCGATCATCGTCATATCATPA 922
|||||
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
|||||
923 GACCAGGAAGTTTATGAGATAATTTGGTCTTTGGTGATATATCAG 972
|||||
251 PheArgTyrThrLeuLeuCystTyrAlaAlaCysGluAlaProLeuIleLe 267
|||||
973 TTCAGATACACACTCCTTTGCTATGCTGCATCGAGGCGCCCACTAATCCT 1022
|||||
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
|||||
1023 TGAATTGGGAGGATATATTATGGAGAAATTCATGTTGTTGTTGAACG 1072
|||||
284 sPTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
|||||
1073 ATTGCATGCCAGCCTTGGCAGTCCCTTCTGCTGCAAAATATAGACCA 1122
|||||
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
|||||

```



1123	TACGGTGTTCACAGATTCCCGCAGCACCCCTTGTATACATAATTTAGC	1172	617	gGlyMetThrLysAspHisThrTrpAspHisAla.ProSerSerThrSer	633
317	aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp	334	2073	AGCATGACGAAAGACCATACGTGGGACCATGCCCGCAGCAGTACGAGC	2122
1173	ACATCAGGGTGTGAGCGCTGCAAGTACATATCTCTGATCTGGGATTGCCAC	1222	634	ArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG1	650
334	roGluTyrTyrGlyAlaLeuGluTrpValPheProGluTyrTrpAlaArgArg	350	2123	AGATCTTCGAATGGCGCTTCGTGGACCAACCCCTACGTCTATGTAGACGGG	2172
1223	CTGAATGTTATGGAGCTTTAGAAATGGGTATTTCCAGAAATGGGCAAGGAGG	1272	650	yLeuGlyArgSerLysCysGlnSerProSerAlaLeuLysThrSerSerS	667
351	HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa	367	2173	ACTGGGAGGTCCAGCGCGGTCTCTTCAGCTCTGAAGACATGTTCTCT	2222
1273	CATGCCCTTGACAGGGGTGAGCGAGTTAACTTTTGAAGAGCAGCAGT	1322	667	erSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAla	683
367	lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT	384	2223	CATCTTCCCGCGCCGGAAGGATACCCCTGTACATTGCGTTGCTCTGCT	2272
1323	GACAGCAGATCGAATTTGACCGGTGACGTCAGGTTATTTCATGGGAGGTCA	1372	684	ThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgTh	700
384	hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys	400	2273	ACAGTAGAGTCGCAATGCGCTGCTTGCT.TGGTCCCGCGGTTCGAGAGT	2321
1373	CNACTGCTGAAGTGGACAGCGCCCTCATGAGCTCTTAAGCTCCGAAAA	1422	700	rTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnL	717
401	SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr	417	2322	AGATGACGGGTGCTGCTCGCGGGTGACAGCTTCGGGTGGATGACAGT	2371
1423	ACTGTATTGAATGGAATTTAAATGGAATTGACATTAAATGATTGGAACC	1472	717	euGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAla	733
417	oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG	434	2372	TACAGTTTGGGGAATTAAGGAAGGATGTCTGCAGGATGGTTAAACAGCA	2421
1473	CACCACAGACAAGTGTCTCCCTCATCANTATTCTGCGTACCTCTCTG	1522	734	LysHisHisSerAspGlySerLeuSerValArgValThrAlaGluIleAr	750
434	lYsAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal	450	2422	AAGCACCTCAGATGCGACCTCTCTGTCGGTGTACAGCTGAATCAG	2471
1523	GAAGGCCAAATGTAAGCTGAATTTGCAGAGGAGCTGGGTTTACCTGTA	1572	750	gAsnGlnLeuValThrLeu	756
451	ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnly	467	2472	AAACCACTGGTGACTCTT	2490
1573	AGGGAGGATGTTCCCTGTGATTGGCTTTATTGGAAGACTGGATTACCAAA	1622	seq_name:	gb_pl:AF091803	
467	sGlyLeuAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV	484	seq_documentation_block:	2591 bp	mrna
1623	AGGCATTGATCTCATTAAATAGCCATTCACAGAGCTCATGAGGAGGAGC	1672	LOCUS	AF091803	14-SEP-1999
484	alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet	500	DEFINITION	Triticum aestivum starch synthase I mRNA, complete cds.	
1673	TGCAGTTTGTCTGCTGTGATCTGGGATCCCAATTTTGAAGCTGGATG	1722	ACCESSION	AF091803	
501	ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValClyPh	517	VERSION	AF091803.1	GI:5880465
1723	AGATCTACCGAGTCGACTTACAAGGATAAATCCCGTGGATGGTTGGATT	1772	KEYWORDS	bread wheat.	
517	eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM	534	SOURCE	Triticum aestivum	
1773	TAGTGTTCAGTTTCCACAGATAACTCAGGTTGCGATATATTGTTAA	1822	ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.	
534	etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln	550	REFERENCE	1 (bases 1 to 2591)	
1823	TGCCATCCAGTTGAACCTTTGGTCTTAATCAGCTATATGCTATGCAAA	1872	AUTHORS	Li,Z., Rahman,S., Kosar-Hashemi,B., Mouille,G., Appels,R. and Morell,M.K.	
551	TyrGlyThrValProValValHisGlyThrGlyLeuArgAspThrVa	567	TITLE	Cloning and characterization of a gene encoding wheat starch synthase I	
1873	TATGGTACAGTTCTCTGATGTAAGTGGGCTCCGAGACACAGT	1922	JOURNAL	Theor. Appl. Genet. 98, 1208-1216 (1999)	
567	lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA	584	REFERENCE	2 (bases 1 to 2591)	
1923	CGAGACCTTCAACCTTTTGGTCAAAAGGAGAGGAGGTACAGGTTGGG	1972	AUTHORS	Li,Z., Rahman,S., Kosar-Hashemi,B., Mouille,G., Appels,R. and Morell,M.K.	
584	laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla	600	TITLE	Direct Submission	
1973	CGTCTCACCCTTAACCTCGCACAGATGTTGTGGCATTTGCCAACCGCG	2022	JOURNAL	Submitted (10-SEP-1998) CSIRO Division of Plant Industry, GPO Box 1600, Canberra, ACT 2601, Australia	
601	MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr	617	FEATURES	Location/Qualifiers	
2023	ATGTCGACATTCAGGGAGCACAGCCGCTCCTGGGAGGGGCTCATGAAGCG	2072	source	1. .2591	
				/organism="Triticum aestivum"	
				/db_xref="taxon:4565"	
				/chromosome="7"	
				/tissue_type="endosperm"	
				256..2199	
				/codon_start=1	
				/product="starch synthase I"	
				/protein_id="AAC54661.1"	
				/db_xref="GI:5880466"	
				CDS	



```

/translaction="MAATGVGAGCLAPSVLRADPATAARASACVVRARLRRLARGRY
VAELSRGCPAARPAQOOOLAPLPVGLFAPPPAPQAPSPAPQPPDAGVGLAPDL
LLEGTAEDSIDIIIVAAEQSESEIDANEQPOAKVTRSIIVFTVEAAAPYAKSGGLGDV
CGSLPIALAARGHRVMVMPRIYLNSSDKNYAKALYTKHKIKPCFGGSHVTFHFHEY
RDNVDMFVDPHSPYHRPGSLYDNGFAGDQNRFTLLCYAAACEAPLILELGGYIYGO
NCMEVNDWHSALVPVLLAAKIRPYGVYRDSRSTLVTHLAHOGLEPASTYDPLGLPP
EMYGALWEFFPEWARRHALDKGEAVNFKGAVTVADRIVTVSOGYSWEVTTABGGGGL
NELSSRSKSVLNGIVNGIDINDWNPYTDKCLPHYSVDDLSGRKAKKAELOKEGLPVP
REDVPLIGFTIGRLDYOKGIDLIKWAPELAREDOVQFVHLGGDPIFEGBMRSTESSYK
DKRGMWGFSPVPSHRITACGDLILMPREFEPCGLNOLYAMQYGTVPVYHGTGGLRDT
VETENPFCAGGEBGTGNWAFSLTVDKMLWALRTAMSTFREHKPSWEGLMKRGMTKDRT
WDHAAREQIEFWAFVFDQPYVM"
BASE COUNT      567 a      672 c      743 g      609 t
ORIGIN

```

```

alignment_scores:
  Quality: 3930.00      Length: 757
  Ratio: 5.247          Gaps: 1
  Percent Similarity: 98.943      Percent Identity: 98.415
alignment_block:
US-09-674-824-2 x AF091803
Align seg 1/1 to: AF091803 from: 1 to: 2591
1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17
|||||
256 ATGGCGGGGACGGCGGTGGGGCGGGTGCCTCGCCCCCGCGCTCCGCT 305
17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
|||||
306 GCGCCCGCATCGCGGACGGCGGGCGGGCGGTCCGCTCGCTCGCTCGCG 355
34 laArgLeuArgArgLeuAlaArgGlyArgTyValAlaGluLeuSerArg 50
|||||
356 CGCGCTCGCGCGCTTGGCGGGGGCGGTACGTGGCGGACTCGGCGAG 405
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
|||||
406 GAGGCGCGCGCGCGCGCGCGCGCGCAGCAGCAACTGGCGCGCGCGCT 455
67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84
|||||
456 CGTGCCAGGCTTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 505
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
|||||
506 CCGCGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 555
101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117
|||||
556 GACCTCTGCTCGAAGGAGTGTGTAGGATTCATCGACAGCATATTGT 605
117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
|||||
606 GGCTGCAAGTGACAGGATTTCTGAGATCATGGATGCGAATGACCAACTC 655
134 lnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
|||||
656 AAGCTAAGTTACACGTAGCATCGTGTGTGTGACTGGTGAAGCTGCTCT 705
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
|||||
706 TATGCAAAAGTCAGGGGGGTGGGAGATGTTGTGGTTCGTACCAATTCG 755
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyLeuA 184
|||||
756 TCTGTGCTGCTGCGTGGTCCGCGGTGTGTGTGTGTGTGTGTGTGTGT 805
184 snGlySerSerAspLysAsnTyAlaLysAlaLeuTyThrAlaLysHis 200
|||||
806 ATGGGTCTCTGTATAAAACTATGCAAGGCATTATACACTGGGAGGCAC 855

```

```

201 lleLysIleProCysPheGlySerHisGluValThrPhePheHisG1 217
|||||
856 ATTAAGATTCCATGCTTTGGGGGATCACATGAAGTGACCTTTTTCATGA 905
217 uTyArgAspAsnValAspTrpValPheValAspHisProSerTyHisA 234
|||||
906 GTATAGACACAACGTCGATTGGGTGTTTTCGATCATCCGTCATATCAT 955
234 rgProGlySerLeuTyGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
|||||
956 GACCAAGAAAGTTATATGGAGATAATTTTGTGCTTTTGTGTGATAATCAG 1005
251 PheArgTyThrLeuLeuCysTyAlaAlaCysGluAlaProLeuIleLe 267
|||||
1006 TTCAGATACACACTCCCTTTCCTATGCTGCATGCGAGGCCCCCACTAATCCT 1055
267 uGluLeuGlyGlyTyTrIleTyGlyGlnAsnCysMetPheValValAsnA 284
|||||
1056 TGAATTGGGAGGATATATTTATGACAGAATTGCATGTTTGTGTGAACG 1105
284 sPTpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyArgPro 300
|||||
1106 ATGGCATGCCAGCTTGTGCCAGTCTCTTCTGCTGCAAAATATAGACCA 1155
301 TyrGlyValTyArgAspSerArgSerThrLeuValIleHisAsnLeuA1 317
|||||
1156 TAGCGTGTTTACAGAGATTCCCGCAGCACCTTGTGTTATACATAATTTAGC 1205
317 aHisGlnGlyValGluProAlaSerThrTyProAspLeuGlyLeuProP 334
|||||
1206 ACATCAGGCTCTGGAGCCTGCAAGTACATCTCTGATCTGGGATTGCCAC 1255
334 roGluTrpTyGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
|||||
1256 CTGAATGGTATGAGCTTTAGAAATGGGTATTTCCAGATGGCGAGGAGG 1305
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
|||||
1306 CATGCCCTTGACAAGGGTGAGGAGTAACTTTTGAAGGAGCAGTCTGT 1355
367 lThrAlaAspArgIleValThrValSerGlnGlyTySerTrpGluValT 384
|||||
1356 GACAGCAGATCGAATTGTGACCGTCAGTCAGGCTTATTCATGGGAGGTCA 1405
384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
|||||
1406 CAACCTGCTGAAGGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCCGAAA 1455
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
|||||
1456 AGTGATTGAATGGAAATTTGAATGGAATTGACATTAAATGATTGGAACCC 1505
417 oThrThrAspLysCysLeuProHisHisTySerValAspAspLeuSerG 434
|||||
1506 CACCACAGACAAGTGTCTCCCTCATCATATTCTGTGATCACCTCTCTG 1555
434 lYlAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
|||||
1556 GAAAGGCCAAATGTAACCTGAATTCGAAAGGAGCTGGGTTTACCTGTA 1605
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyGlnLy 467
|||||
1606 AGGAGGAGTGTCTCTGATTGGCTTTTATGGAAGACTGGATTACACGAA 1655
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
|||||
1656 AGGATTGATCTCAATAAATGGCCATTTCCAGAGCTCATGAGGAGGAGC 1705
484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTyTrpMet 500
|||||
1706 TGCAGTTTGTCTGATCTGGGATCTGGGATCCCAATTTTGAAGGCTGGATG 1755

```



ORGANISM Aegilops tauschii  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Aegilops.  
REFERENCE 1 (bases 1 to 2662)  
AUTHORS Li,Z., Morell,M. and Rahman,S. in plants  
TITLE Regulation of gene expression  
JOURNAL Patent: WO 9914314-A 11 25-MAR-1999;  
GOODMAN FIELDER LTD (AU) ; LI ZHONGYI (AU) ; MORELL MATTHEW (AU) ;  
RAHMAN SADEOUR (AU) ; UNIV AUSTRALIAN (AU) ; COMM SCIENT IND RES  
ORG (AU) ; GROUPE LIMAGRAIN PACIFIC PTY L (AU)  
FEATURES  
source Location/Qualifiers  
1..2662  
/organism="Aegilops tauschii"  
/db\_xref="taxon:37682"  
/tissue\_type="ENDOSPERM"  
misc\_feature 1..2651  
BASE COUNT 592 a 677 c 754 g 639 t  
ORIGIN  
alignment\_scores:  
Quality: 3930.00 Length: 757  
Ratio: 5.247 Gaps: 1  
Percent Similarity: 98.943 Percent Identity: 98.415  
alignment\_block:  
US-09-674-824-2 x AX031272 ..  
Align seg 1/1 to: AX031272 from: 1 to: 2662  
1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17  
|||||  
257 ATGGCGCGACGGCGCTCGGGCGGGGTGGCTCGCCCCACAGCTCGCGCT 306  
17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34  
|||||  
307 CGCGCGCGATCCGCGACGGCGGGCGGCGTCCGCTCGCTCGCTCGCG 356  
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50  
|||||  
357 CGCGGCTCGGCGCTTGGCGGGGCGGCTACGTTGCCGAGCTCAGCAGG 406  
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67  
|||||  
407 GAGGCT 456  
67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84  
|||||  
457 CGTGCAGGCTTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 506  
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100  
|||||  
507 CCCCACACGACCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 556  
101 AspLeuLeuLeuGluGlyLeuAlaGluAspSerIleAspSerIleIleVa 117  
|||||  
557 GACCTCTCTGCTGAAGGATTTGCTGAGGATTTCCATCGACAGCATATGT 606  
117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnPro 134  
|||||  
607 GGCTCAAGTGACAGGATTTCTGAGATCATGATCGGAATGAGCAACCTC 656  
134 lAlaAlaValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150  
|||||  
657 AAGCTAAAGTTACGTAGCATCTGTTGTGACTGGTGAAGCTGCTCTCT 706  
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167  
|||||  
707 TATGCAAAAGTCAGGGGGGCTGGGAGATGTTTGTGTTGTTACCAATTGC 756  
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184

501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517  
|||||  
1756 AGATCTACCGAGTCGAGTACAAGGATAAATCCGTGGATGGGTGGATT 1805  
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534  
|||||  
1806 TAGTGTTCCTCAGTTCCCCACAGAACTACGAGTTCGATATATTGTAA 1855  
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMeGln 550  
|||||  
1856 TGCCATCCAGGTTGAACCTTTGTGTCTTAATCAGCTATATGCTATCAA 1905  
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567  
|||||  
1906 TATGTACAGTTCCTGTGTAGTTCATGGAACTGGGGCCCTCCGAGACACAGT 1955  
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584  
|||||  
1956 CGAGACCTTCAACCCCTTTTGTGTGCAAAAGGAGAGGAGGGGTACAGGTGG 2005  
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600  
|||||  
2006 CGTCTCACCGCTAACCCCTGGACAAGATGTTGTGGGCATTCGGAACCCG 2055  
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617  
|||||  
2056 ATGTGACATTCAGGAGACACAAAGCCCTCTGGGAGGGGCTCATGAAGCG 2105  
617 gGlyMetThrLysAspHisThrTrpAspHisAla.ProSerSerThrSer 633  
|||||  
2106 AGGCATGACGAAGACATACATGTTGGGACCATGCCCGCGAGCATGACGAGC 2155  
634 ArgSerSerGlyProSerTrpThrAsnProThrSerCysArgArgGl 650  
|||||  
2156 AGATCTTCGAATGGGCTTCGTGGACCAACCTACGTCTGCTAGACGGG 2205  
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667  
|||||  
2206 ACTGGGAGTCTGAAGCGCGGCTCTCTTGTAGCTCTGAAGACATGTTCT 2255  
667 erSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAla 683  
|||||  
2256 CATCTTTCGCGCGCGCGGGAAGGATACCCCTGTACATGCTGCTCTGCT 2305  
684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaClySerArgTh 700  
|||||  
2306 ACAGTAGAGTCGCAATGGCGCTCTGCT.TGGTCCGCGCGGTTCGAGAGT 2354  
700 rTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyArgGlnL 717  
: |||||  
2355 AGATGACGCTGTGCTGCTGCGCGGTGACAGCTTCGGGTGGATGACAGT 2404  
717 euGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAla 733  
|||||  
2405 TACAGTTTTGGGGAATGAAGGAGGATGCTGCACAGATGGTTAACAGCA 2454  
734 LysHisSerAspGlySerLeuSerValArgValThrAlaGluLeuAr 750  
|||||  
2455 AAGCACCACTCAGATGGCAGCCTCTCTGCTCCGTGTTACAGCTGAATCAG 2504  
750 gAsnGlnLeuValThrLeu 756  
|||||  
2505 AAACCAACTGGTGACTCTT 2523  
seq\_name: gb\_pat:AX031272  
seq\_documentation\_block:  
LOCUS AX031272 2662 bp DNA PAT 20-SEP-2000  
DEFINITION Sequence 11 from Patent WO9914314.  
ACCESSION AX031272  
VERSION AX031272.1 GI:10278604  
KEYWORDS Aegilops tauschii.  
SOURCE



|||||  
757 TCTTGCTGCTCGTGGTCAACCGTGTGATGGTTGTAATGCCAAGACTTGA 806  
|||||  
184 snglySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200  
|||||  
807 ATGGGTCTCTGATAAAACTATGCAAGGCATTATACACTGGGAAGCAC 856  
|||||  
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG1 217  
|||||  
857 ATTAAGATTCCATGCTTTGGGGATCACATGAAGTGACCTTTTTCATGA 906  
|||||  
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234  
|||||  
907 GTATAGAGACAACGTCGATTGGGTGTTGTGTCATCATCCGTCATATCATA 956  
|||||  
234 rGProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250  
|||||  
957 GACCAGGAAGTTATATGGAGATAAATTTGGTGCTTTTGGTGATAATCAG 1006  
|||||  
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267  
|||||  
1007 TTCAGATACACACTCCTTCGTATGTCATGCGAGGCGCCCACTAATCCT 1056  
|||||  
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284  
|||||  
1057 TGAATTGGGAGGATATATTATGGACAGAATTGCATGCTTTGTTGTGAACG 1106  
|||||  
284 spTrpHisAlaSerLeuValProValLeuLeuAlaLysTyrArgPro 300  
|||||  
1107 ATGGGATGCCACCTTGTGCCAGTCCTTCTGTGTCGCAAAATATAGACCA 1156  
|||||  
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317  
|||||  
1157 TACGGTGTTTACAGAGATTCCCGCAGACCCCTTGTATACATAATTAGC 1206  
|||||  
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334  
|||||  
1207 ACATCAGGGTCTGGAGCCTGCAAGTACATATCCTGATCTGGGATTGCCAC 1256  
|||||  
334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350  
|||||  
1257 CTGAATGGTATGAGCTTTAGATGGGTATTCGAAGTGGCAAGGAGG 1306  
|||||  
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValva 367  
|||||  
1307 CATGCCCTTGACAAGGCTGAGGCAGTTAACTTTTGAAGGACACGTCGT 1356  
|||||  
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384  
|||||  
1357 GACAGCAGATCGAATTTGTGACCGTCAGTCAGGGTTATTTCATGGGAGGTCA 1406  
|||||  
384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400  
|||||  
1407 CAACTGCTGAAGTGCACAGGGCTCAATGAGCTCTTAAGTCCCGAATAA 1456  
|||||  
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417  
|||||  
1457 AGTGTATTGAATGGAATTTGTAATGGAATTGACATTAATGATTGGAACCC 1506  
|||||  
417 oThrThrAspLysCysLeuProHisIleTyrSerValAspLeuSerG 434  
|||||  
1507 CACCACAGACAGTGTCCTCATCATATTCTGTCGATGACCTCTCTG 1556  
|||||  
434 lYlysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 1606  
|||||  
1557 GAAAGCCAAAATGAAGCTGAATTCGAGAAGGAGCTGGGTTTACCTGTA 1656  
|||||  
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467  
|||||  
1607 AGGGAGGATGTTCTCTGATGGCTTTATTGGAAGACTGGATTACCAGAA 1656  
|||||  
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484  
|||||

1657 AGCATTGATCTCATTTAAATGGCCATTCCAGAGCTCATGAGGGAGCAGC 1706  
|||||  
484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500  
|||||  
1707 TGCAGTTTGTTCATGCTTGGATCTGGGATCCAATTTTGAAGGCTGGATG 1756  
|||||  
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517  
|||||  
1757 AGATCTACCAGAGTCGAGTTACAGGATAAATCCGTGGATGGGTGGATT 1806  
|||||  
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534  
|||||  
1807 TAGTGTTCAGTTTCCACAGAATACTGCAGGTTGCGATATATGTTAA 1856  
|||||  
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550  
|||||  
1857 TGCATCTCCAGGTTTGAACCTTGTGGTCTTAATCAGCTATATGCTATGCAA 1906  
|||||  
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567  
|||||  
1907 TATGGTACAGTTCTCTGATGTTTCATGGAACCTGGGGGCTCCGAGACACAGT 1956  
|||||  
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpA 584  
|||||  
1957 CGAGACCTTCAACCCCTTTTGGTGCARAAAGCAGAGGAGGTACAGGTTGG 2006  
|||||  
584 lPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600  
|||||  
2007 CGTCTCACCGCTTAACCGTGGACAAGATGTTGGGCATTCGGAACCGCG 2056  
|||||  
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617  
|||||  
2057 ATCTCACATTCAGGAGCACAAGCCGCTCTGGAGGGGCTCATGAAGCG 2106  
|||||  
617 gGlyMetThrLysAspHisThrTrpAspHisAla.ProSerSerThrSer 633  
|||||  
2107 AGGCATGACGAAAGACCATACGTGGGACCATGCCCGCAGCAGTACGAGC 2156  
|||||  
634 ArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG1 650  
|||||  
2157 AGATCTCGAATGGGCTTCGTGGACCAACCCCTACGTCATGTAGACGGG 2206  
|||||  
650 yLeuGlyArgSerLysCysGlySerProSerAlaLeuLysThrSerSerS 667  
|||||  
2207 ACTGGGAGGTCGAGCGCGGCTCTCTTTCAGCTCTGAACACATGTTCT 2256  
|||||  
667 erSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAla 683  
|||||  
2257 CATCCTTCGCGCGCGGAGGATACCCCTGTACATTGCGTTGCTCTGCT 2306  
|||||  
684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgTh 700  
|||||  
2307 ACAGTACAGTCGCANTGCGCTCTGCT.TGGTCCGCGCGGTTTCGAGAGT 2355  
|||||  
700 rTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnL 717  
: |||||  
2356 AGATGACGCGCTGTGCTGCTCGCGCGGTGACAGCTTCGGGTGGATGACAGT 2405  
|||||  
717 euGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAla 733  
|||||  
2406 TACAGTTTTGGGAATAAGGAAGGATGTGCTGCAGGATGGTTAACAGCA 2455  
|||||  
734 LysHisHisSerAspGlySerLeuSerValArgValThrAlaGluIleAr 750  
|||||  
2456 AAGCACCACTCAGATGGCAGCCTCTCTGTCCGTGTACAGCTGAATTCAG 2505  
|||||  
750 asnGlnLeuValThrLeu 756  
|||||  
2506 AAACCAACTGGTGACTCTT 2524  
|||||

seq\_name: gb\_pat:A93350

seq\_documentation\_block:



LOCUS A93350 2239 bp DNA PAT 22-JAN-2000  
DEFINITION Sequence 1 from Patent WO9745545.  
ACCESSION A93350  
VERSION A93350.1 GI:6741617  
KEYWORDS bread wheat.  
SOURCE Triticum aestivum  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.  
REFERENCE 1 (bases 1 to 2239)  
AUTHORS Block, M. and Loerz, H.  
TITLE NUCLEIC ACID MOLECULES ENCODING ENZYMES FROM WHEAT WHICH ARE  
INVOLVED IN STARCH SYNTHESIS  
JOURNAL Patent: WO 9745545-A 1 04-DEC-1997;  
HOECHST SCHERING AGREVO GMBH (DE); BLOCK MARTINA (DE)  
FEATURES  
source  
1. .2239  
/organism="Triticum aestivum"  
/strain="CV. FLORIDA"  
/db\_xref="taxon:4565"  
/clone="TASSS"  
/haplotype="CA. 21 D CARYOPSES"  
/clone\_lib="CDNA LIBRARY IN PBLUESCRIPT SK (-)"  
<3. .2018  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB69543.1"  
/db\_xref="GI:6741618"  
/translation="TQPPLDAGVGLADPLLEGLAEDSIDSIIVAAEQSEIMDA  
NEPOAKVTRSVFTVGEAAYAKSGGLDVGCSLPALAAARHVRVMPYRLNGSS  
DKNYAKALYAKHIKIPCRGGSHVTFHEYRDNDVDFVDPSPYHRPGSLYDNFGA  
FDGNFRVTLCLYAAACEAPLLELGGYIVGQNCMFVNDWHSALVPLLAARYPVGV  
YRDSRTLVHNLAHQGVPEASTYDGLGPPPEWYGALEWVPEWARRHALDKGEAVNF  
LGAVYADRIYVSGYSWETTABGGGGLNELSSRSKVLNGIVNGIDINDWPTT  
DKLPHYSVDLSGKAKKAEQLKELGVPREDVPLIGFIRGLDTQKIDILKMAIP  
ELMRDVPYGLSGSDGPIPEGMRSSTESSYKDKFRWGFVSPVSHRITAGCDILMP  
SRPEPCGLNOLYAMQYGVVYVHGTGLRDVETENPFCAKEEGTGWAFSPLTDKM  
LWALRTAMFTGHEKPSWGLMKRMTKDHWPASSTSRSSGSPSWNPTSCRRGL  
GRSKCPSALATSSSRPGEGYPCPLCPATVESQACILWFAFSRTYDGCAGAAV  
TASGGRLQFWIRKGCAGAGWLTAKHHSGLSLVRVTAIRNQLVTL"  
BASE COUNT 611 a 448 c 590 g 590 t  
ORIGIN  
alignment\_scores:  
Quality: 3612.00 Length: 671  
Ratio: 5.383 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-674-824-2 x A93350 ..  
Align seg 1/1 to: A93350 from: 1 to: 2239  
86 ThrGlnProLeuProAspAlaGlyValGlyGluLeuAlaProAspLe 102  
3 ACGAGCGCCCGCTGCCGACGCCGGCTGGGGAACCTCGCGCCGACCT 52  
102 uLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleValAla 119  
53 CCTGCTCAAGGGATTGCTGAGGATTCATCGACAGCATATATTGGCGTG 102  
119 laSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProGlnAla 135  
103 CAAGTGACGAGATTCTGAGATCATGATGCGATGACACACCTCAAGCT 152  
136 LysValThrArgSerIleValPheValThrGlyGluAlaAlaProTyrAl 152  
153 AAAGTTACACGTAGCATCGTGTGACTGCTGGAAGCTGCTCCTTATGC 202  
152 aLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuA 169  
|||||

203 AAAGTCAGGGGGTGGAGATGTTTGTGGTTCGTACCAATTCCTCTTG 252  
169 laAlaArgGlyHisArgValMetValValMetProArgTyrLeuAsnGly 185  
|||||  
253 CTGCTCTGGTCAACCGAGTGATGCTTGTAAATGCCAAGATACTTAAATGGG 302  
186 SerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLe 202  
|||||  
303 TCCTCTGATAAAACTATGCAAGGCAATTATACACTGCGAAGCACATTAA 352  
202 siLeProCysPheGlyGlySerHisGluValThrPhePheHisGluTyr 219  
|||||  
353 GATTCACATGCTTTGGGGATCAGATGAGTGACCTTTTTCATGAGTATA 402  
219 rgAspAsnValAspTrpValPheValAspHisProSerTyrHisArgPro 235  
|||||  
403 GAGACAACGTCGATTGGGTGTTTTCGATCATCCGTCATATACACAGACA 452  
236 GlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheAr 252  
|||||  
453 GGAAGTTTATATGAGATAAATTTGGTGCTTTTGGTGATAATCAGTTTCA 502  
252 gTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGlu 269  
|||||  
503 ATACACACTCCTTTGCTATGCTGATCGAGGCCCGCCACTAATCTTGAAT 552  
269 euGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrp 285  
|||||  
553 TGGGAGGATATATTTATGACAGAAATTCATGTTTGTGTAACGATTGG 602  
286 HisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgProTyrGl 302  
|||||  
603 CATGCCACCTTGTGCCAGTCCTTCTTCTGCAAAATATAGACCATACGG 652  
302 yValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHis 319  
|||||  
653 TGTTTACAGAGATTCCCGCAGACACCTTTGTTATACATAATTTAGCACAT 702  
319 lnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProGlu 335  
|||||  
703 AGGTGTGGAGCCCTGCAAGTACATATCCTGATCTGGGATTCCTCTGAA 752  
336 TrpTyrGlyAlaLeuGluTyrValPheProGluTyrAlaArgArgHisAl 352  
|||||  
753 TGGTATGGAGCTTTAGAAATGGGTATTTCCAGAATGGCAAGGAGGCATGC 802  
352 aLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValValThr 369  
|||||  
803 CCTTGACAAGGGTGAGGAGTAACTTTTGAAGGAGCAGTTGTGACAG 852  
369 laAspArgIleValThrValSerGlnGlyTyrSerTrpGluValThrThr 385  
|||||  
853 CAGATCGGATTGTGACCGTCAGTCAGGTTTATTCATGGGAGTCACAACT 902  
386 AlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerVa 402  
|||||  
903 GCTGAAGGTGGACAGGGCTCAATGAGCTCTTAAGCTCCCGAAAAAGTGT 952  
402 lLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThr 419  
|||||  
953 ATTGAATGGAATGTAAATGGAATTTGACATTAATGATTTGGAACCCCA 1002  
419 hrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerGlyLys 435  
|||||  
1003 CAGACAAGTGTCTCCCTCATCATTTATCTGTCGATGACCTCTCTGGAAG 1052  
436 AlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGl 452  
|||||  
1053 GCCAAATGTAAGCTGAAATTCAGAAGGAGTTGGGTTTACCTGTAAAGGA 1102  
452 uAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyI 469  
|||||  
1103 GGATGTTCTCTGATTTGGCTTTTATTTGGAGACTGATTTACCAGAAAGCA 1152



```

469 leAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGln 485
|||||
1153 TTGATCTCATTAATGATGCCATCCAGAGCTCATGAGGAGGAGCGTGCAA 1202
486 PheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgse 502
|||||
1203 TTTGTGCATGCTTGGATCTGGGATCCAAATTTTGAAGGCTGGATCAGATC 1252
502 rThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerV 519
|||||
1253 TACCCAGTTCGAGTTACAAGGATAAATCCGTGGATGGGTGGATTTAGTG 1302
519 alProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetPro 535
|||||
1303 TTCCAGTTTCCACACAGAATACTGCAGGTCCGATATATTGTTAATGCCA 1352
536 SerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrG 552
|||||
1353 TCGAGATTTGAACCTTGGCGTCTTAATCAGCTATATGCTATGCAATATGG 1402
552 yThrValProValValHisGlyThrGlyGlyLeuArgAspThrValGluT 569
|||||
1403 TACAGTTCTCTAGTTTCATGAACCTGGGGCTCCGAGACACAGTCGAGA 1452
569 hrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpAlaPhe 585
|||||
1453 CTTTCAACCCCTTTGGTGCAAAAGGAGGAGGAGGTACAGGTGGCGTTC 1502
586 SerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSe 602
|||||
1503 TCACCGCTAACCGTGGACAGATGTTGTGGCATTTGCGAACCCGATGTC 1552
602 rThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyM 619
|||||
1553 GACATTCAGGAGACACAGCGCTCTGGAGGGCTCATGAAGCGAGGCA 1602
619 etThrLysAspHisThrTrpAspHisAlaProSerSerThrSerArgSer 635
|||||
1603 TGACGAAGACCATACCTGGGACCATGCCCGGAGCAGTAGCAGCATCT 1652
636 SerSerGlyProSerTrpThrAsnProThrSerCysArgArgGlyLeuG 652
|||||
1653 TCGAGTGGGCTCTCGTGGACCAACCTACGTCATGTAGACGGGACTGGG 1702
652 YArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSerSerP 669
|||||
1703 GAGGTCCAGTGCAGTCTCCTTCAGCTCTGAAGACATCTCTTCATCCT 1752
669 heArgGlyProGluGlyTyrProCysThrLeuArgCysProAlaThrVal 685
|||||
1753 TCCGCGCGCCGGAAGGATACCCCTGTACATTGCGTTGCTCTCTACAGTA 1802
686 GluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrTyrAs 702
|||||
1803 GAGTCCGAATGCGCCCTGCTGCTTTGTTGCGCGGTTCGAGAACATATCA 1852
702 pGlyCysAlaAlaAlaValThrAlaSerGlyArgGlnLeuGlnP 719
|||||
1853 CGGCTGCTGCTGCGCGGTGACAGCTTCGGGTGGACACAGTATACAGT 1902
719 heTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAlaLysHis 735
|||||
1903 TTTGGGAATAAGGAAGGATGTGTCAGAGTGTTAACACCAAGGACAC 1952
736 HisSerAspGlySerLeuSerValArgValThrAlaGluIleArgAsnG 752
|||||
1953 CACTCAGTGGGAGGCTCTCTGCTCGGTGTACAGCTGAATCAGAAACCA 2002
752 nLeuValThrLeu 756
|||||
2003 ACTGGTGACTCTT 2015

```

seq\_name: gb\_pl:TAU48227

seq\_documentation\_block: 2055 bp mRNA PLN 12-JUN-1996

LOCUS TAU48227 Triticum aestivum soluble starch synthase mRNA, partial cds.

DEFINITION Triticum aestivum soluble starch synthase mRNA, partial cds.

ACCESSION U48227 GI:1373149

VERSION U48227.1

KEYWORDS wheat.

SOURCE Triticum aestivum

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 2055)

AUTHORS Block, M., Loerz, H. and Luetticke, S.

TITLE Direct Submission

JOURNAL Submitted (01-FEB-1996) Martina Block, University of Hamburg,

Institute of General Botany, Centre of Applied Molecular Biology,

AMP II, Ohnhorstst. 18, Hamburg, 22609, Germany

COMMENT On Jun 12, 1996 this sequence version replaced gi:1335887.

FEATURES Location/Qualifiers

1..2055

/organism="Triticum aestivum"

/cultivar="T.a., Florida"

/db\_xref="taxon:4565"

/tissue\_type="endosperm of kernels"

<1..1834

/codon\_start=2

/product="soluble starch synthase"

/protein\_id="AAB02197.1"

/db\_xref="GI:1373150"

/translation="EAAPYAKSGGLGVCSLPIAARHVRVMVMPYRLNGSSDKN

YAKALYAKHIKIPCFGSGSEVTFHEYRDNVDVDFVDFHPSHRPGSLYCDNFAGFD

NQRYTLIYACAEAPLILELGYIYQNCMFVYNDWHASLVPLLAARYPYGVYRD

SRSTLVHNLHAGQVPEASTYDGLPPEWYGALEWVPEWARHRLADKGEAVNFKG

AVVTADRIYVSQGYSWEVTTAEGGOLNELLSSRKSVLNGIVNGIDINDWPTDKC

LPHYYSVDLLSGKAKAELOKELGLPVREDVPLIGFGLDYQKGLDLIKNAIPELM

REDVQVMLGSGDPIFECHMRSTESSYKDKFRGWGFSVPVSHRITAGCDLILMPSRF

EPGQLNQLYAMQYGVVPHKGTDRDVTETNPFKAGEEGTGWAFSLPTVDKMLWA

LRYANSTFRHFKHSGWELKMGMTKDHDPSPSSSRSSGSPNTNPTSCRRGLGRS

KCESPLAKITSSSSFRGEGYPTCLRCPATVSCQACLLWFAFSRTYDGC AAAAATAS

GGROLQFGIRKGCAAGWLTAKHSDGSLSVRVTAIRNQLVTL"

BASE COUNT 570 a 399 c 534 g 552 t

ORIGIN

alignment\_scores:

Quality: 3313.00 Length: 610

Ratio: 5.431 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-674-824-2 x TAU48227

Align seg 1/1 to: TAU48227 from: 1 to: 2055

147 GluAlaAlaProTyrAlaLysSerGlyGlyLeuGlyAspValCysGlySe 163

2 GAAGCTGCTCTTATGCAAGTCAAGGGGGTTGGAGATGTTGTGTTTC 51

163 rLeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValValMetP 180

52 GTTACCAATTTGCTCTTCTGCTGCTGCTACCGAGTGTGTTGTTATGC 101

180 roArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyr 196

102 CAAGATCTTAAATGGGTCTCTGATAAAACTATGCAAGAGCATTTATAC 151

197 ThrAlaLysHisIleLysIleProCysPheGlyGlySerHisGluValTh 213

152 ACTGCGAAGCACATTAAGATTCCATGCTTTGGGGGATCAGATGAGTGAC 201

213 rPhePheHisGluTyrArgAspAsnValAspTrpValPheValAspHisP 230



```
|||||
202 CTTTTTTCATGAGTATAGACAACTCGATTGGGTGTTGTCGATCAATC 251
230 roSerTyrHisArgProGlySerLeuTyrGlyAspAsnPheGlyAlaPhe 246
252 CGTCATATCACAGACCAGGAAGTTTATATGGAGATAATTTGGTGCTTT 301
247 GlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAl 263
302 GCTGATAATCAGTTCAGATACACACTCCTTTGCTATGCTGCATCGAGGC 351
263 aProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetP 280
352 CCCACTAATCCTGAAATGGGAGGATATATTATGGACAAATGGATGT 401
280 heValValAsnAspTrpHisAlaSerLeuValProValLeuLeuAlaAla 296
402 TTGTTGTGAAGATTGGCATGCCAGCCTTGCCAGTCTTCTTGCTGCA 451
297 LysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLeuValI 313
452 AAATATAGACCATACGGTGTGTACAGAGATTCCCGCAGCACCCCTTG 501
313 eHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrProAspL 330
502 ACATAATTTAGACATACGGTGTGTACAGAGATTCCCGCAGCACCCCTTG 551
330 euGlyLeuProGluTyrPyrGlyAlaLeuGluTyrValPheProGlu 346
552 TGGGATTCCTCCCTGAATGGTATGAGCTTTAGAAATGGGTATTTCCAGAA 601
347 TrpAlaArgArgHisAlaLeuAspGlyGlyGluAlaValAsnPheLeuLy 363
602 TGGGCAAGGAGCATGCCCTTGACAAGGCTGAGGAGTTAACTTTTGA 651
363 sGlyAlaValValThrAlaAspArgIleValThrValSerGlnGlyTyr 380
652 AGGACACGTTGTGACAGCAGATCGGATGTGACCGTCAGTCAGGGTATT 701
380 erTrpGluValThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeu 396
702 CATGGAGGTCAACAACCTGCTGAAGGTGACAGGCGCTCAATGAGCTCTTA 751
397 SerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAspIleAs 413
752 AGCTCCCGAAAAAGTATTGAATGGAATTTGAATGGAATTGACATTA 801
413 nAspTrpAsnProThrThrAspLysCysLeuProHisHisTyrSerVala 430
802 TGATTTGGAACCCACACAGCAAGTGTCTCCCTCATCATATTATCTGTCG 851
430 sAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGlnLysGluLeu 446
852 ATGACCTCTCTGGAAGGCCAATGTAAAGCTGAATTCAGAAAGGAGTTG 901
447 GlyLeuProValArgGluAspValProLeuIleGlyPheIleGlyArgLe 463
902 GGTTTACCTGTAAGGAGGAGTGTCTCTGATTTGCTTTATTTGGAAGACT 951
463 uAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIleProGluLeuM 480
952 GGATTTACCGAAGAGGAGTGTCTCTCATTAATAATGGCCATTCAGAGCTCA 1001
480 eArgGluAspValGlnPheValMetLeuGlySerGlyAspProIlePhe 496
1002 TGAGGAGGAGCGTCAATTTGTCATGCTTGGATCTGGGATCCAATTTT 1051
497 GluGlyTrpMetArgSerThrGluSerTyrLysAspLysPheArgG 513
1052 GAAGGCTGGATGAGATCTACCGAGTCTGAGTTACAAGATAAATTCCTGG 1101
513 yTrpValGlyPheSerValProValSerHisArgIleThrAlaGlyCysA 530
```

```
1102 ATGGGTTGGATTTAGTGTTCAGTTTCCACAGAAATAACTGCAGGTTGCG 1151
530 spIleLeuLeuMetProSerArgPheGluProCysGlyLeuAsnGlnLeu 546
1152 ATATATTGTTAATGCCATCGAGATTTGAACCTTGGGTCCTTAATCAGCTA 1201
547 TyrAlaMetGlnTyrGlyThrValProValValHisGlyThrGlyGlyLe 563
1202 TATGCTATGCAATATGGTACAGTTCTCTGTAGTTTCATGGAACCTGGG 1251
563 uArgAspThrValGluThrPheAsnProPheGlyAlaLysGlyGluGluG 580
1252 CCGAGACACAGTCGAGACCTTCAACCTTTTGGTCGCAAAAGAGAGAGG 1301
580 LysThrGlyTrpAlaPheSerProLeuThrValAspLysMetLeuTrpAla 596
1302 GTACAGGTTGGGCTTCTCACCGCTAACCGTGAGCAAGATGTTGTGGCA 1351
597 LeuArgThrAlaMetSerThrPheArgGluHisLysLysProSerTrpGlu 613
1352 TTGCGAACCGCATGTCGACATTCAGGGAGCACAAAGCCGTCCTGGGAGG 1401
613 yLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHisAlaProS 630
1402 GCTCATGAAGCGAGGCATGACGAAAGCCATACGTGGGACCATGCCCGCA 1451
630 exSerThrSerArgSerSerGlyProSerTrpThrAsnProThrSer 646
1452 GCAGTACGAGCAGATCTTCGAGTGGGCTTCGTGGACCAACCCCTACGTCA 1501
647 CysArgArgGlyLeuGlyArgSerLysCysGluSerProSerAlaLeuLy 663
1502 TGTAGACGGGACTGGGGAGGTCCAAGTCGAGTCGAGTCTCCTTCAGCTG 1551
663 sThrSerSerSerPheArgGlyProGluGlyTyrProCysThrLeuAla 680
1552 GACATCTCTTCATCTTCCGGCGCCGGAAGGATACCCCTGTACATTGC 1601
680 rgCysProAlaThrValGluSerGlnCysAlaCysLeuLeuTrpPheAla 696
1602 GTTGTCTCTGCTACAGTAGAGTCGCAATGGCCCTGCTTGTGTTGGTTCG 1651
697 GlySerArgThrTyrAspGlyCysAlaAlaAlaValThrAlaSerG 713
1652 GGTTCGAGAACATATGACGCTGTGCTGCTGCGCGGTGACAGCTTCGGG 1701
713 yGlyArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyT 730
1702 TGGACGACAGTTACAGTTTGGGGAATAAGGAAGGATGTGCTGCAGGAT 1751
730 rpLeuThrAlaLysHisHisSerAspGlySerLeuSerValArgValThr 746
1752 GGTAAACAGCAAGCACCACTCAGATGGCAGCCTCTCTGTCGCTGTACA 1801
747 AlaGluIleArgAsnGlnLeuValThrLeu 756
1802 GCTCAATCAGAAACCAACTGGTGACTCTT 1831
```

seq\_name: gb\_pl:AF165890

seq\_documentation\_block: 2585 bp mRNA PLN 17-AUG-1999  
LOCUS AF165890  
DEFINITION Oryza sativa subsp. japonica soluble starch synthase mRNA, complete cds.

ACCESSION AF165890  
VERSION AF165890.1 GI:5734102

KEYWORDS  
SOURCE Oryza sativa subsp. japonica.  
ORGANISM Oryza sativa subsp. japonica.

REFERENCE  
1 (bases 1 to 2585)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.



```

· AUTHORS      Junwang,X. and Zhen, Z.
TITLE         Direct Submission
JOURNAL       Submitted (06-JUL-1999) Group 601, Genetics Institute of CAS, DaTun
FEATURES      Road, Beijing, Beijing 100101, China
SOURCE        Location/Qualifiers
              1..2595
                /organism="Oryza sativa subsp. japonica"
                /cultivar="Nanjing37"
                /sub_species="japonica"
                /db_xref="taxon:39947"
              139..2019
                /codon_start=1
                /product="soluble starch synthase"
                /protein_id="AAB49850.1"
                /db_xref="GI:5734103"
                /translation="MATAAGMGTAACLVAPQVRGRRRLQVRVRRRCVAELSRDGGG
AQRPLAPALPKQVLPTFLVPTSTPAPTQSPAPATPPPLPDSGVGETEPDLEGIT
ESDIDKTIIVASEQSEIMDVKEQAQKVTRSVFVTGASPKYAKGGGLGDCVGSILPI
ALALGRHYVMNPNRYMGALCNFNANFYETKHKIKPCFGGEHYETFEHYRDSVDW
VFVDYTHRPNLYGDNFGAFCGNQFRYTLACYACEAPLILELGYIYGQKCMFW
NWHASLVPVLLAAKYPYGVYRDARSVLVHNLHAGVEPASTYPDGLGPPPEYMGAL
EWFPPEWARRHLDKGEAVNFTLKGAVVIADRVITVSQGYSEWVTTAEGGGGLNELLGS
KSVLGVNIVGIDINDMNSTFKFLPHYHSVDLDSGKAKCAELOKELGLPIPDVPLI
IGFGRDLQKGTDLIKLAIPLDMRDNIOQVYMLGSGDPCFEGWMRSTESGYRDKFRG
VGFSPVSHRITAGCDLILAMPSPFEPCFGLNAMYGYVPVYVHGTTGGRLDRIYENENP
FAEKGEQGTGWAFFSLTIETKNAVGIADGNFDIQGTQVLLGGSNEARHVRKLYMGPCRL
tv"
BASE COUNT   660 a   568 c   706 g   651 t

```

IV.			
BASE COUNT	660 a	568 c	706 g
651 t			

```

alignment_scores:
  Quality: 2829.50
  Ratio: 4.527
  Percent Similarity: 86.088
  length: 726
  Gaps: 18
  Percent Identity: 76.446

alignment_block:
  US-09-674-824-2 x AF165890
  ..
  Align seg 1/1 to: AF165890 from: 1 to: 2585

```

```

2 AlaAlaThrGlyValcIyAlaGlyCysLeu...AlaProSerValArgLe 17
151 GCGGGATCGGATCGGGCGCGTGCCTGTGTGCCGCGAGGTGAG... 198
17 uArgAlaAspProAlaThraAlaAlaSerAlaCysvalValArgA 34
199 |||||...CCGGGGAGG.....AGST 211
34 laArgLeuArgArgLeuAlaArgGlyArgTyrrValAlaGluLeuSerArg 50
212 TCGGGGTCCACGGCGGTGGCGAGG...CGTGTCGTGGCGGAGCTGACGAG 258
51 GluCylProAlaAlaArgProAlacInclnclnleuAlaPro..... 65
259 GACGTGGGTGGCGG.....CAGGCCCGCTGGCACCGCGCC 296
66 .....ProLeuValProGlyPheLeuAlapro.....Prop 76
297 GCTGTGTGAAGCAGCCGGTCTCGCGGACCTCTCGTSCCGACGTCGACGC 346
76 roProAlaProAlaGlnSer.....ProAlaProThrGlnProProLeu 90
347 CACCCGCGCCACGACGATCGCGCGCGCGCGGACCCGCGCGCGTTG 396
91 ProAspAlaGlyValGlyGluLeuAlaProAspLeuLeuLeuGluGly 107
397 CGCGACTCCGCGTGGGGAGATCGACGCCGAT.....CTAGAAGGTCT 440
107 eAlaGluAspSerlleAspSerlleIleValAlaAlaSerGluInAspS 124
441 CACAGNAGATTCCATCGACAACAAATTTTGGCTGTAGTGACGAGGAT 490
```

124	erGlulLeuMetAspAlaAsnGluInProGlnAlaLysValThrArgSer	140
491	CTGAGATCATGGATGCTGAAGGACCAAGCTCAAGCTAAAGTAACACGCAGC	540
141	IleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeu	157
541	GTGTGCTTTGTACCCGGTGAAGCTTCTCCTTATGCAAAATCAGGTGGAGCT	590
157	uGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHisA	174
591	AGGAGATGTTTGTGGTTCACTGCCAATTGCTCTCTCTCTCGTGGTCATC	640
174	rgValMetValValMetProArgTyrLeuAsnGlySerSerAspLysAsn	190
641	GTGTGATGTTGTATGCCGATACATGACGGGGCTTGAACAAAAAT	690
191	TyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheG	207
691	TTTGCAACGCCATTTTACACTGAGAAGCACATTAAAGATTCCATGCTTTGG	740
207	YGLYSerHisGluValThrPhePheHisGluTyrArgAspAsnValAsp	224
741	CGGAGAACATGAAGTTACTTTTTTCACGAGTATAGGAGATTCTGTGATT	790
224	rpValPheValAspHisProSerTyrHisArgProGlySerLeuTyrGly	240
791	GGTGTGTTTGTATCATCCCTCATATCATACACTGGAAATTTGATGGA	840
241	AspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCy	257
841	GATAAATTTGTGCTTTTGGCGATAATCAGTTCAGATACACACTCCGTG	890
257	styrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIle	274
891	CTATCGCGGTGTGAAGCCCATTAATTTCTTGAATGGAGGATATATCT	940
274	YrGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuVal	290
941	ATGGACAGAAATGATGTTTGTGTGAATGATTGGCATGCCAGTCCGTGTG	990
291	ProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSe	307
991	CCAGTCTCTTCTGTCGCAAAATATAGACCATATCGTGTTTACAGGGATGC	1040
307	rArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProA	324
1041	CCGCAGTGTCTTGTCTATACATAATCTAGCACATCAGGGTGTGGAGCCTG	1090
324	LaSerThrTyrProAspLeuGlyLeuProGluTyrTyrGlyAlaLeu	340
1091	CCAGTACATATCTTGACTGGGATGGCCACTGAATGGTATGGAGCATTA	1140
341	GluTrpValPheProGluTrpAlaArgHisAlaLeuAspLysGlyG	357
1141	GAATGGGTGTGTTCCAGAGTGGCGAAGCGGCATGCCCTTCACAAGGTGA	1190
357	uAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleVal	374
1191	GGCAGTCAATTTTTAAAGGCGCAGTTGTGGACAGCAGATCGGAATTTGTA	1240
374	hrValSerGlnGlyTyrSerTrpGluValThrThrAlaGluGlyGlyGln	390
1241	CTGTACGCCAGGGGTATTCTATGGAGGTCACTGCTGAGGTGGGCATA	1290
391	GlyLeuAsnGlnLeuLeuSerSerArgLysSerValLeuAsnGlyIleVal	407
1291	GGCCTCAATGAGCTCTTAAGCTCCCGAAGAGTGTATGGAATGGAATCTG	1340
407	lAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuP	424
1341	AAATGGAAATTGACATTAATGATTGGAAACCCATCCACAGACAGTTTCTCC	1390
424	roHisHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysAla	440



```

seq_documentation_block: 2533 bp RNA PAT
LOCUS E06904
DEFINITION cDNA encoding soluble starch synthase.
ACCESSION E06904
VERSION E06904.1 GI:2175060
KEYWORDS JP 1994070779-A/1.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheop-
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriartoideae; Oryzaceae; Oryza.
1 (bases 1 to 2533)
Baba,T. and Shimada,H.
TITLE SOLUBLE RICE STARCH SYNTHETASE GENE AND ITS USE
JOURNAL Patent: JP 1994070779-A 1 15-MAR-1994;
MITSUI GIYOUSAI SHOKUBUTSU BIO KENKYUSHO:KK
COMMENT OS Oryza sativa
PN JP 1994070779-A/1
PD 15-MAR-1994
PF 07-JUL-1992 JP 1992179947
PI BABA TADASHI, SHIMADA HIROAKI
PC C12N15/52,A01H5/00;
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue_type=ripening seed;
FH *source: tissue_type=ripening seed;
FH Key Location/Qualifiers
FT 5'UTR 1..113
FT CDS 114..1994
FT /product='Soluble starch synthase' FT
FT transit_peptide 114..452
FT 3'UTR 1992..2533.
FEATURES
source
1..2533
/organism='Oryza sativa'
/db_xref='taxon:4530'
630 a 550 c 701 g 652 t
BASE COUNT
ORIGIN
alignment_scores:
Quality: 2821.50 Length: 726
Ratio: 4.522 Gaps: 18
Percent Similarity: 85.950 Percent Identity: 76.309
alignment_block:
US-09-674-824-2 x E06904 ..
Align seg 1/1 to: E06904 from: 1 to: 2533
2 AlaAlaThrGlyValGlyAlaGlyCysLeu...AlaProSerValArgLe 17
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
126 GCGGGGATGGGATCGGGCGGCGTGCCTGGTGGCGCGGAGGTGAGG.. 173
17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
||||| ||||| ||||| ||||| ||||| ||||| |||||
174 .....CCGGGGAGG.....AGGT 186
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
||||| ||||| ||||| ||||| ||||| ||||| |||||
187 TGGCGGCTCAGCGGGTGGCGGAG...CGGTGGTGGCGAGCTGAGCAGG 233
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaPro..... 65
||||| ||||| ||||| ||||| ||||| ||||| |||||
234 GACGGTGGGTGGCGG.....CAGCGCGCGTGGCACCAGCGGCC 271
66 .....ProLeuValProGlyPheLeuAlaPro.....Prop 76
||||| ||||| ||||| ||||| ||||| ||||| |||||
272 GCTGTGTGAAGACCGCGTCTCGCCACCTTCTCGTGGCGAGCTGACGC 321
||||| ||||| ||||| ||||| ||||| ||||| |||||

```



|||||.....  
322 CACC CGCCGCGCAGTGC GGGCGCGGCCGACCGCGCGCGGTG 371

91 ProAspAlaGlyValGlyGluLeuAlaProAspLeuLeuGluGlyLyl 107  
|||:::|||:::  
372 CCGCACTCCGGTG GGGGAGATCGAGCCCGAT .....CTAGAAGTCT 415

107 eAlaGluAaspSerIleAspSerIleValAlaLaSerGluInAsps 124  
:::::::::::  
416 CACGAAGATTTCATCGACAACAAATTTTGGTGTACTGACGAGGACT 465

124 erGlulMetAspAlaAsnGluInProGlnAlaLysValThrArgSer 140  
:::::::::::  
466 CTGAGATCATGATGTGAAGGAGCAAGCTCAAGCTTAAGATAACACGCAGC 515

141 lleValPheValThrGlyGluAlaLaProTyrAlaLysSerGlyGlyLe 157  
:::::::::::  
516 GTTGCTTTGTAAACCGGTGAAGCTTCTCTTATGCAAAATCAGGTGGACT 565

157 uGlyAspValCysGlySerLeuProIleAlaLeuAlaAargGlyHisA 174  
:::::::::::  
566 AGGAGATGTTGTGGTCTACTCCCAATTGCTCTGCTCTCGTGGTGCATC 615

174 rGValMetValValMetProArgTyrLeuAsnGlySerSerAspLysAsn 190  
:::::::::::  
616 GTGTGATGTTGTAATGCCGAGATACATGAAGGGCGCTTGAACAAAAT 665

191 TyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheGl 207  
:::::::::::  
666 ITTGCAACGCAATTTTACACTGAGAAGCACATTAAGATTCCATGCTTTGG 715

207 yGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValAspt 224  
:::::::::::  
716 CGGACAACATGAAGTTACTTTTTTCAGCAGTAGAGGATTCGTGTGATT 765

224 tpValPheValAspHisProSerTyrHisArgProGlySerLeuTyrgly 240  
:::::::::::  
766 GGGTGTTTGTGATCATCCCCTCATATCATACACCTGGAAATTTGTATGGA 815

241 AspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCy 257  
:::::::::::  
816 GATAATTTTGGTCTTTTGGCGATAATCAGTTCAGATACACACTCCGTGTG 865

257 sTyrAlalaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrileT 274  
:::::::::::  
866 CTATCGCGCGTGTGAAGCCCCAATAATTCTTGAACCTGGAGGATATATCT 915

274 yrGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuVal 290  
:::::::::::  
916 ATGCACAGAAATGCATGTTGTTGTAATGATTGGCATGCCAGTCCTTGTGTG 965

291 ProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgSpse 307  
:::::::::::  
966 CCAGTCTCTTGTGTGCAAAATATAGACATATGCTGTTTACAGGGATGC 1015

307 rArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProA 324  
:::::::::::  
1016 CCGCAGTGTCTTGTCTATACATAATCTAGCACATCAGGGGTCTGGAGCCTG 1065

324 laSerThrTyrProAspLeuGlyLeuProGlnTyrTrpTyrGlyAlaLeu 340  
:::::::::::  
1066 CCAGTACATATCTGACCTGGGATTGCCACCTGAATGSTATGAGCATTA 1115

341 GlutrpValPheProGluTTrpAlaArgArgHisAlaLeuAspLysGlyGl 357  
:::::::::::  
1116 GAATGGGTGTTTCCAGAGTGGGCAAGCGCGCATGCCCTTCACAAGGGTGA 1165

357 uAlaValAsnPheLeuLysGlyAlaValAlaThrAlaAspArgIleValT 374  
:::::::::::  
1166 GGCAGTCAAATTTTTTAAAAAGCGCAGTGTGACAGCAGATCGAATGTGTGA 1215

374 hrValSerGlnGlyTyrSerTrpGluValThrAlaGluGlyGlyGln 390  
:::::::::::

1216 CTGTCA GCCAGGGGTATTTCATGGGAGGTCACAACTGCTGAAGGTGGCAA 1265

391 GlyLeuAsnGlnLeuLeuSerSerArgLysSerValLeuAsnGlyIleVal 407

1266 GGCCTCAATGAGCTCTTAAGCTCCCGGAAGAGTGTATTGAATGGAATGT 1315

407 IAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuP 424

1316 AAATGGAAATGACATTAATGATTTGAACCCATCCACACACAAGTTTCTCC 1365

424 roHisHisTrpSerValAspAspLeuSerGlyLysAlaLysCysLysAla 440

1366 CTTATCATTTACTGTGTGATGACCTGTCCGGAAGGCCAGTGTAAAGCT 1415

441 GluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuII 457

1416 GAATTCGAGAAGGAGCTGGTTTACCTATAAAGGCCGATGTGCCTCTGAT 1465

457 eGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysM 474

1466 TGCCTTTATTGGAAGATTGGCACTATCAAAAAGGCATTTCATCTAATTAAC 1515

474 etaLaiIleProGluLeuMetArgGluAspValGlnPheValMetLeuGly 490

1516 TTGCCATTCAGATCTCATCGGGACAATATTCATTCGTCACTGCTTGGGA 1565

491 SerGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSerty 507

1566 TCTGGTACCCAGTTTTTCAAGATGGATGAGATCCACAGAAATCAGGGTA 1615

507 rLysAspLysPheArgGlyTrpValGlyPheSerValProValSerHisA 524

1616 CAGGNTAAATTCGTGGATGGTTGGATTAGTTGTTCCTAGTTCCACC 1665

524 rgIleThrAlaGlyCysAspLLeuLeuMetProSerArgPheGluPro 540

1666 GAATAACTGCAGTTGCGCATATATGTGTGATGCCATCCAGATTTCGAACCT 1715

541 CysGlyLeuAsnGlnLeuTyrAlaMetClnTyrClyThrValProValVa 557

1716 TGTGGCTCAATCAGCTATATGCTATGCATATGTTACAGTGGCTGTGT 1765

557 IHisGlyThrGlyLeuArgAspThrValGluThrPheAsnProPheG 574

1766 TCATGGAACTGGAGCCCTCAGAGATACAGTGGAGAAATTTTAAACCCGTTC 1815

574 lYAlaLysClyClnClnClyThrGlyTrpAlaPheSerProLeuThrVal 590

1816 CTGAGAAAGGAGACAGGGTACAGGGTGGGCATCTCGCCATTAACCAT 1865

591 AspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluH 607

1866 GAAAAAATCTGTGGCATTCGGATGGCAATTCGACATACAGGCAAC 1915

607 isLysProSerTrpClnGlyLeuMetLysArgGlyMetThrLysAspHis 623

1916 ACAAGTCCTCTGGGAGGGTCTAATGAAGCGAGCATGTCAAACGCAATT 1965

624 ThrTrpAspHisAlaProSerThrSerArgSerSerSerGlyProS 640

1966 ACATGGGACCATCGCCCTCACAGTATGACAGATCTTCGAATGGCCCTT 2015

640 erTrpThrAsnProThrSerCysArgArgGlyLeuGlyArgSerLys 655

2016 CATGGATCAACCATATGTCATGTAATGGATTGTAAGGAGGACGAGCAATT 2065

656 CysGluSerProSerAlaLeuLysThrS 665

2066 TCTCGGAGGACCTCAATCTCTCTGCTTCATGAGCGGAATGAAACCTT 2115

665 erSerSerSerPheArgGlyPr 672

2116 TGTACATCATAGTAAAGGAGCAACAGTTATGCAAGCTTCGAAACATCAC 2165



672 oGluGlyTyrProCysThrLeuArgCysProAlaThrValGluSerGlnC 689  
 ::::::::::::::::::::|  
 2166 TCAAGGTACCCCTGTAGG.....CCGTACTTGGCCA..... 2199

689 ysAlaCysLeuLeuTrrPhe 695

2200 .....ATATGGTTC 2208

seq\_name: gb\_pl:RICSSS

seq\_documentation\_block: 2533 bp mRNA PLN 01-FEB-2000  
 LOCUS RICSSS Rice mRNA for soluble starch synthase, complete cds.  
 DEFINITION D16202.1 GI:450484  
 VERSION soluble starch synthase.  
 KEYWORDS Oryza sativa Immature seed, cDNA to mRNA, clones RS[1, 2, 3 and 4].  
 SOURCE Oryza sativa

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 2533)

Baba,T., Nishihara,M., Mizuno,K., Kawasaki,T., Shimada,H.,

Kobayashi,E., Ohnishi,S., Tanaka,K. and Arai,Y.

Identification, cDNA cloning, and gene expression of soluble starch

synthase in rice (Oryza sativa L.) immature seeds

Plant Physiol. 103 (2), 565-573 (1993)

94302151

2 (bases 1 to 2533)

Baba,T.

Direct Submission

Submitted (06-MAY-1993) to the DDBJ/EMBL/GenBank databases. Tadashi

Baba University of Tsukuba, Institute of Applied Biochemistry;

Tennohdai 1-1-1, Tsukuba Science City, Ibaraki 305, Japan

(Tel:298-53-6632, Fax:298-53-6632)

Submitted (06-May-1993) to DDBJ by:

Tadashi Baba

Institute of Applied Biochemistry

University of Tsukuba

Tennohdai 1-1-1

Tsukuba Science City

Ibaraki 305

Japan

Phone: 0298-53-6632.

Fax: 0298-53-6632.

Location/Qualifiers

1. .2533

/organism="Oryza sativa"

/db\_xref="taxon:4530"

/dev\_stage="Immature seed"

1. .113

transit\_peptide 114..452

CDS 114..1994

/codon\_start=1

/product="soluble starch synthase precursor"

/protein\_id="BAA03739.1"

/db\_xref="GI:450485"

translation="MATAAGMIGAAACLVAPVRPGRRLRLQRVRRRCVAELSRDGGG

AHGLAPAPLVKQPLTFLPTSTPAPTQSPAPPTPPLPDSSGVEIPELEGIT

EDSIDKTFIVASEGESEIMDVKEQAKVTRSVVFTGEASPKSGGLGDVCGSLPI

ALALGRHVWVPMRYMNGALKNFANFYTERHIKIPCEGGEHEVTFEYRDSVDM

VFDHPHSRPNGLYGNFGAFGNQFRTLLCYACEAPLILELGGYIYGQCMFVV

NDMHASLPVLLAAKYRPYGVYDRSVLVTNHLAHQGVPEASTYDPLGLPPEMYGAL

EMVPEWARHALDKGAVNFKGAVVTADRIVTVSOGYSWEVTTAEGGGLNELLS

RKVLNGTVNGIDINDNPNSTDKFLPVHVSDDLSGKAKCAEOLKELGLP IRPDVPL

IGFTGRDLYQKIDILIKLIPDLMRDNIQVWMLSGSDPFGEGHWRSTESGYRDKFRGM

VGFSPVSHRITACGDILLMPSEPEPGLNQLYAMQYGTVPVHGTGGLRDTVENFNP

FAERGEQGTGWAFLPTIEKNAVGIDNFDIQGTQVLLGGSNEARHVKRLYMGPCRL

TV"

mat\_peptide 453..1991

/product="soluble starch synthase"

3'UTR 1995..2533

polya\_signal 2487..2492  
 BASE COUNT 630 a 550 c 701 g 652 t  
 ORIGIN

alignment\_scores:

Quality: 2821.50 Length: 726

Ratio: 4.522 Gaps: 18

Percent Similarity: 85.950 Percent Identity: 76.309

alignment\_block:

US-09-674-824-2 x RICSSS ..

Align seg 1/1 to: RICSSS from: 1 to: 2533

2 AlaAlaThrGlyValGlyAlaGlyCysLeu...AlaProSerValArgLe 17  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 126 CGGGGATGGGATCGGGCGGGCGGTGCTGCTGGTGGCGCGGAGTGAGG... 173  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 174 .....CCGGGGAGG.....AGGT 186  
 34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 187 TGC GGCTCCAGCGGGTCCGGAGG...CGGTGGTGGCGGAGCTGAGCAGG 233  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaPro..... 65  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 234 GACGGTGGTGGCG.....CACGGCGCGTGGCACCGCGGCC 271  
 66 .....ProLeuValProGlyPheLeuAlaPro.....Prop 76  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 272 GCTGGTGAAGCAGCGGTCTCGCGACCTTCCTCGTCCGCGAGTCGACGC 321  
 76 roProAlaProAlaGlnSer.....ProAlaProThrGlnProProLeu 90  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 322 CACCGCGCGCCACGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGTTG 371  
 91 ProAspAlaGlyValGlyGluLeuAlaProAspLeuLeuGluGlyI 107  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 372 CCGGACTCCGCGGTGGGGAGATCGAGCCGAT.....CTAGAAGGTCT 415  
 107 eAlaGluAspSerIleAspSerIleIleValAlaAlaSerGluGlnAsps 124  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 416 CACAGAGATTCCTCATGACAAACAATTTTGTGGCTAGTAGGAGGAGT 465  
 124 erGluIleMetAspAlaAsnGluGlnProGlnAlaLysValThrArgSer 140  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 466 CTGAGATCATGTGATGTGAAGGAGCAGCTCAAGCTAAGTAACACGACGC 515  
 141 lleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLe 157  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 516 GTTGCTTTGTAAACCGGTGAAGCTTCTCTTATGCAAGTCAGGTGGACT 565  
 157 uGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHisa 174  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 566 AGGAGATGTTGTGGTTTCACTGCCAAATTTGCTTCTCTTCTGCTGTCATC 615  
 174 rgValMetValValMetProArgTyrLeuAsnGlySerSerAspLysasn 190  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 616 GTGTGATGTTGTATGCGGAGATACATGACGCGGCGCTTGAACAAAAAT 665  
 191 TyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheG 207  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 666 TTTCACAAACGCAATTTTACACTGAGAACACACATTAAGATTCCTGCTTGG 715  
 207 yGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValAspt 224  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 716 CGGAGAACATGAAGTTACTTTTTTTCACGAGTATAGGATTCGTGTGATT 765  
 224 rpValPheValAspHisProSerTyrHisArgProGlySerLeuTyrGly 240  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||



```
766 GGGTGTGTTTGTGATCATCCCTCATATCATAGACCTGGAAATTTGTATGA 815
241 AspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCy 257
816 GATAATTTTGGTGTGTTTGGCGATAATCAGTTCAGATACACACTCCTGTG 865
257 sTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleT 274
866 CTATGCGGCGTGAAGCCCAATTAATTCCTGAAGTGGAGGATATATCT 915
274 yGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuVal 290
916 ATGSGACAAATGCAATGCTTTGTTGTAATGATGGCATGCCAGTCTTGTG 965
291 ProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSe 307
966 CCAGTCTCTTGTGTCGCAAAATATAGACCATATGTTGTTTACAGGGATGC 1015
307 rArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProA 324
1016 CCGCAGTCTTCTGTGCATACATAATCTAGCACATCAGGGTGTGAGCGCTG 1065
324 laSerThrTyrProAspLeuGlyLeuProProGluTrpTyrGlyAlaLeu 340
1066 CCAGTACATATCCAGCATGCGGATGGCCACCTGAATGATGGAGCATTA 1115
341 GluTrpValPheProGluTrpAlaArgHisAlaLeuAspLysGlyG1 357
1116 GAATGGGTGTTTCCAGAGTGGCAGGCGGCTGCCCTTGACAAGGGTGA 1165
357 uAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValT 374
1166 GGCAGTCAATTTTAAAGCGCAGTGTGACAGCAGATCGAATTTGTGA 1215
374 hrValSerGlnGlyTyrSerTrpGluValThrAlaGluGlyGlyGln 390
1216 CTGTCAGCCAGGGTATTCATGGGAGGTGCACACTGCTGAAGTGGGCAA 1265
391 GlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleVa 407
1266 GGCCTCAATGAGCTCTTAAAGCTCCGGAAGAGTGTATTGAATGAATGT 1315
407 laSnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuP 424
1316 AAATGGAATTGACATTAATGATTGGAACCATCCACAGCAAGTTTCTCC 1365
424 roHisHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysAla 440
1366 CTTATCATTTATCTGTTGATGACCTGTCGGAAGGCCAAGCTGTAAAGCT 1415
441 GluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuIl 457
1416 GAATTCGAGAAGGAGCTGGGTTTACCTATTAAGGCCCGCATGTGCTCTGAT 1465
457 eGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysM 474
1466 TGCGTTTATTTGGAGATGTGACTATCAAAAGGCATGATCTAATTAAC 1515
474 etAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuGly 490
1516 TTGCCATTCAGATCTCATCGGGACAATATTCAATTCGTCATGCTTGA 1565
491 SerGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSerTy 507
1566 TCTGGTGACCCAGGTTTGAAGGATGGATGAGATCCACAGAATCAGGGTA 1615
507 rLysAspLysPheArgGlyTrpValGlyPheSerValProValSerHisA 524
1616 CAGGGATAAATTTTCGTGGATGGGTTGGAATTAGTGTTCAGTTTCCACC 1665
524 rgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluPro 540
1666 GAATAACTGCAGGTTCGCATATATTGTTGATGCCATCCAGATTCGAACCT 1715
```

```
541 CysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValVa 557
1716 TGTGSCCTCAATCAGCTATATGCTATATCAATATGTTACAGTGCCTGTGT 1765
557 lHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheG 574
1766 TCATGGAACCTGGAGGCTCAGAGATACAGTGGAGAATTTTAACCCGTTTG 1815
574 yAlaLaLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrVal 590
1816 CTGGAAGAAGGAGAGCAGGGTACAGGGTGGGCAATTCGCCCACTAACCAT 1865
591 AspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluH 607
1866 GAAAAAATGCTGTGGGCATTCGGGATGGCAATTTGCACATACAGGGAAC 1915
607 isLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHis 623
1916 ACAAGTCTCTTGGGAGGGTCTAATGAAGCGAGGCATGTCAAGCGACTTT 1965
624 ThrTrpAspHisAlaProSerSerThrSerArgSerSerSerGlyProS 640
1966 ACATGGGACCATCCCGCTCACAGTATGACAGATCTTCAATGGGCCCTT 2015
640 erTrpThrAsnProThrSerCysArgArgGlyLeuGlyArgSerLys.. 655
2016 CATGGATCAACCATATCTCATGTAAATGGATTTGAAGAGCAGCGCAATT 2065
656 CysGluSerPro.....SerAlaLeuLysThrS 665
2066 TCTCCAGGACCTCAATCTCTCTGCTTCATGAGCGGAATGAAAACTT 2115
665 erSerSerPheArgGly.....Pr 672
2116 TGTACACTACATGGAAGGGAACCACTTATGCAAGTTCCAAACGATCAC 2165
672 oGluGlyTyrProCysThrLeuArgCysProAlaThrValGluSerGlnC 689
2166 TCAAGGTTACCTTTGAGG.....CCTGCTACTTGGCCA..... 2199
689 ysAlaCysLeuLeuTrpPhe 695
2200 .....ATATGGTTC 2208
seq_name: gb_pat:A93359
seq_documentation_block:
LOCUS A93359 2383 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9744472.
ACCESSION A93359
VERSION A93359.1 GI:6741623
KEYWORDS
SOURCE
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 2383)
AUTHORS Kossmann, J. and Froberg, C.
TITLE NUCLEIC ACID MOLECULES CODING SOLUBLE MAIZE STARCH SYNTHASES
JOURNAL Patent: WO 9744472-A 1 27-NOV-1997;
KOSSMANN JENS (DE); FROBERG CLAUS (DE)
FEATURES
source
1..2383
/organism="Zea mays"
/db_xref="taxon:4577"
/tissue_type="ENDOSPERM"
<2..1951
/function="STARCH SYNTHESIS"
/codon_start=1
/product="SOLUBLE STARCH SYNTHASE"
/protein_id="CAB69545.1"
```







```

501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
1508 AGATCTACAGAGTCGATCTTTCAAGGATAAAATTCGTGGATGGCTTGGATT 1557
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
1558 TAGTGTTCACAGTTTCCACCAGAAATAACTGCCGGCTGCGATATATTGTAA 1607
534 eProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
1608 TGCATPCCAGATTCCGAACCTTGTGGTCTCAATCAGCTATATGCTATGCAG 1657
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567
1658 TATGGCACAGTTCTCTGTGTCCATCGAACTGGGGGCTTAGAGATACCGT 1707
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584
1708 GGAGAACTTCAACCCCTTTCGGTCGAAATGGAGAGCAGGCTACAGGTGGG 1757
584 lPheSerProLeuThrValAspLysMetLeuThrTrpAlaLeuArgThrAla 600
1758 CATTCGACCCCTAACCCACAGAAACATGTTGTGGACATTGGCAACTGCA 1807
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
1808 ATATCTACATACAGGAACACAAAGTCTCTCGGAAGGCGTAAATGAAGCG 1857
617 gGlyMetThrLysAspHisThrTrpAspHisAlaPro_SerSerThrSer 633
1858 AGGCATGTCAAAGACTTTCATCGTGGACCATCGCGTAACCAATACGAAC 1907
634 ArgSerSerGlyProSerTrpThrAsnProThrSerCysArgArgGly 650
1908 AAATCTCCAGTGGCCCTTCATCGATCGACCTATGTCATGTAAANAAG 1957
650 yLeuGlyArgSerLysCys_GluSerProSerAlaLeuLysThrSerSer 666
1958 A.....CCAAAGTGGTGGTCTCTTGAAGATCATCATGTTTCATCATCC 1998
667 SerSerPheArgGlyProGluGlyTrpPro.....CysThrLeuArgCy 681
1999 TATACTAGCTCAATGATGAAGAAACCCCTGTACATTACATGAAGAGC 2048
681 sProIaThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlyS 698
2049 AGACCGCGTATTGGTCCATTGCTCCAATGCTGCTGCTTTGGCTGCCT.... 2094
698 eArgThrTyrAspGlyCysAlaAlaAlaAlaValThrAlaSerGlyGly 714
2095 .....TGCCTCGATGGACCGGATGCGAGTGGAGGAATCCA 2127
715 ArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAla.....Al 728
2128 GCCGAACACAGTTTTGAAGGATAGGAAGGGAGCGTGGACCACTGACGC 2177
728 aclyTrpLeuThrAlaLysHisHisSerAspGlySerLeuSer 742
2178 AGGCAGCCTCCCGGTGATTCATATGGAACAAGCTGGAGTCAGT 2220
seq_name: gb_pl:AF036891

```

<b>seq_documentation_block:</b>					
LOCUS	AF036891	2991 bp	mRNA	PLN	03-MAY-2001
DEFINITION	Zea mays starch synthase I precursor ( <i>Ssl1</i> ) mRNA, nuclear gene encoding plastid protein, complete cds.				
ACCESSION	AF036891				
VERSION	AF036891.1 GI:2828011				
KEYWORDS	"				
SOURCE	Zea mays.				
ORGANISM	Zea mays				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta Tracheophyta;				

REFERENCE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Zea.
AUTHORS	1 (bases 1 to 2991) Knight, M.E., Harn, C., Lilley, C.E.R., Guan, H., Singletary, G.W., Mu-Forster, C., Wasserman, B.P. and Keeling, P.L.
TITLE	Molecular cloning of starch synthase I from maize (W64) endosperm and expression in <i>Escherichia coli</i>
JOURNAL	Plant J. 14 (5), 613-622 (1998)
MEDLINE	98340555
PUBMED	9675904
REFERENCE	2 (bases 1 to 2991)
AUTHORS	Knight, M.E., Harn, C., Lilley, C.E.R., Guan, H.P., Singletary, G.W., Mu-Forster, C., Wasserman, B.P. and Keeling, P.L.
TITLE	Direct Submission
JOURNAL	Submitted (93-DEC-1997) ExSeed Genetics, 1568 Food Science Building, ISU, Ames, IA 50013, USA
FEATURES	Location/Qualifiers 1...2991

[illegible]

```

alignment_scores:
  Quality: 2776.50      Length: 747
  Ratio: 4.442         Gaps: 11
  Percent Similarity: 83.668      Percent Identity: 73.762

alignment_block:
  US-09-674-824-2 x AF036891      ..

Align seg 1/1 to: AF036891 from: 1 to: 2991

4   ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20
614  TCGGCCCTGGCGCGCGGTCTCTCTC..... 640
20  pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34
641  .CTCGCGCGGCGCGCTGGCGCGCGCGCTGCGGCGACCGGCGCGCGCGC 899
35  ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



690 GCGCGCTCCAGCGCGTGTGCGCCCGCGTGTGCGCGAGCTGAGCAGG 739  
 51 GluGlyProAlaIaAArgProAlaGlnGlnGlnGlnLeuAlaProProLe 67  
 740 GAGGGCGCGCGCGCGCGCGTGCACCCCGCGTGTGCGCGCCCGCT 789  
 67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84  
 790 CGTGGCGCGCTCTCTCGCG...CGCGCGCGAGCCACGGGTGAGCGG 836  
 84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100  
 837 CATCGACGCGCGCGCGTGTGCGCGCGCGCTGCGGGGACCTCGCT... 883  
 101 AspLeuLeuLeuGlyGlyLeuAlaGluAspSerIleAspSerIleIleVa 117  
 884 ...CTCGAACCTGAAGGATTGCTGAAGTTCCATCGATACACAGTAGT 930  
 117 lAlaIaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134  
 931 TCTGCCAAGTGACGAAGATTCTGACATTGTGGTGGAAAGGAGCAAGCTC 980  
 134 laAlaIysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150  
 981 GAGCTAAAGTAACAAAGCATTTGCTTTGTAACCGCGCAAGCTTCTCCT 1030  
 151 TyrAlaIysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167  
 1031 TATGCNAAGTCTGGGGTCTAGACATGTTGTGGTTCTATGCCAGTTGC 1080  
 167 aLeuAlaIaArgGlyHisArgValMetValValMetProArgTyrLeuA 184  
 1081 TCTTGTGCTGCTGTGTCACCGTGTGATGTTCTTAATGCCAGATATTAA 1130  
 184 sngLysSerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200  
 1131 ATGGTACCTCCGATAGAATTATGCAATGCAATTTTACACAGAAACAC 1180  
 201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG 217  
 1181 ATTCGATTCCATGCTTTGGCGGTGAACATGAAGTTACCTTCTTCATGA 1230  
 217 uTyrArgAspAsnValAspIrpValPheValAspHisProSerTyrHisA 234  
 1231 GTATAGAGATTACGTTGACTGGGTGTTTGTGATCATCCCTCATATACA 1280  
 234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250  
 1281 GACCTGGAAATTTATATGAGGATAAGTTTGGTGTCTTGGTGTATATCAG 1330  
 251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267  
 1331 TTCAGATACACACTCTCTTGTCTATGCTGCTGCTGAGGCTCTCTTGTATC 1380  
 267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284  
 1381 TGAATTTGGGAGGATATTTATGGACAGAATTGCAATTTGTGTGCAATG 1430  
 284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300  
 1431 ATTTGGCATGCCATCTAGTGCCAGTCTCTTCTGCTGCAAAATATAGACCA 1480  
 301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317  
 1481 TATGGTGTATTAAGACTCCCGACATCTTGTGTAATACATAATTTAGC 1530  
 317 aHisGlnGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuPro 334  
 1531 ACATCAGGTTGTAGACCTGCAAGCACATATCTGACCTTGGTGGTCCAC 1580  
 334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350  
 1581 CTGAATGGTATGAGCTCTCGAGTGGGTATTCCCTGAATGGCGGAGGAGG 1630

351 HisAlaLeuAspTysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367  
 1631 CATCCCTTGACAGGCTGAGGCAGTTAAATTTTGAAGGTCAGTTGT 1680  
 367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384  
 1681 GACAGCAGATCGAATCGTACTGTCTAGTAAGGTTATTCTGGGAGGTCA 1730  
 384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400  
 1731 CAACTGCTGAAGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCAGAAAG 1780  
 401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417  
 1781 AGTCTATTAAACGGAATTGTAATGGAATTGACATTAAATGATTGGAACCC 1830  
 417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434  
 1831 TGCCACAGACAAATGTATCCCTCTCATTAATCTGTGTGATGACCTCTCTG 1880  
 434 lYlLysAlaLysCysGlyAlaGluLeuGlnLysGluLeuGlyLeuProVal 450  
 1881 GAAAGCCAAATGTAAAGTGCATTGCGAGGAGAGCTGGGTTTACCTATA 1930  
 451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467  
 1931 AGGCTGATGTTCTCTGATTGGCTTTATTGGAAGTTGGATTATCAGAA 1980  
 467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484  
 1981 AGGATTGATCTCAATCACTATCATACCAGATCTCATGCGGGAAGATG 2030  
 484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500  
 2031 TTCATTTGTCATGCTTGGATCTGCTGACCCAGAGCTTGAAGATTGGATG 2080  
 501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517  
 2081 AGATCTACAGATCTCAAGGATAAATTTCTGGATGGGTGGATT 2130  
 517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534  
 2131 TAGTGTCCAGTTTCCACCCAGCAATAACTGCCGCTGCGATATATTGTAA 2180  
 534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550  
 2181 TGCCATCCAGATTCCGACCTTGTGCTCTCAATCAGCTATATGCTATGCG 2230  
 551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567  
 2231 TATGTCAGATTCTCTTGTTCATGCAACTGGGGCCTTAGAGATACCGT 2280  
 567 lGluThrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTyrPa 584  
 2281 GGAGAACTTCAACCTTTTCGGTGAGATGGAGAGCAGGCTACAGGTTGGG 2330  
 584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuAtqThrAl 600  
 2331 CATTCGACCCCTTAACACAGAAACATGTTTGTGGACATTCGCACTGC 2380  
 600 aMetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysA 617  
 2381 AATATCTACATACAGGGAACAACTCTCTCTGGAAGGCTAATGAAGC 2430  
 617 rgGlyMetThrLysAspHisThrTrpAspHisAlaProSerSerThrSe 633  
 2431 GAGGATGTCAAAGACTTCACTGGGACCATGCGGCTGAACAATACGAA 2480  
 633 rArgSerSerGlyProSerTrpThrAsnProThrSerCysArgArg 649  
 2481 CAAATCTTCCAGTGGGCTTTCATCGATCGACCC...GATGTTCAATGGA 2527



650 GlyLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSe 666  
 2528 AAAAAGGACCAAGTGGTGGTCTTCTGAAGATTATCAGTTTCATCATC 2577  
 666 rSerSerPheArgGlyProGluGlyTyrPro.....CysThrLeuArgC 681  
 2578 CTATAGTAAGCTGAAGTGAAGAAAGAAACCCCTGTATACATGGAAGG 2627  
 681 ysProAlaThrValGluSerGlnCysAlaCysLeuLeuTrpPhe..... 695  
 2628 CACACCGGCTATTGGCTCCATGCTCCAAATGCTGCTTTGGCTGCTTGC 2677  
 695 ..... 695  
 2678 CTCGATGGACCGGATGACGTAGGAAATCCAGNCGAAGACAGCTTTGAAG 2727  
 696 .....AlaGlySerArgThrTyrAspGlyCysAlaAlaAlaA 708  
 2728 GATAGGAAGGAGCTGGAAAGC...AGTCACGACGAGGCAAGCCTTCG 2774  
 708 laValThrAlaSerGlyArgGlnLeuGlnPhe 719  
 2775 CCGTTAATTCATATGCAACAACTGAGGTCAGTTT 2809

seq\_name: gb\_pl:AF168786

seq\_documentation\_block:  
 LOCUS AF168786 2592 bp mRNA PLN 03-JAN-2001  
 DEFINITION Sorghum bicolor soluble starch synthase mRNA, complete cds.  
 ACCESSION AF168786  
 VERSION AF168786.2 GI:12019655  
 KEYWORDS  
 SOURCE sorghum.  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 1 (bases 1 to 2592)  
 Hsieh, J.S., Chen, M.R. and Hsing, Y.I.C.  
 Molecular cloning of a Sorghum cDNA encoding the soluble starch  
 synthase SBSS  
 Unpublished  
 2 (bases 1 to 2592)  
 Hsieh, J.S., Chen, M.R. and Hsing, Y.I.C.  
 Direct Submission  
 Submitted (14-JUL-1999) Agronomy, National Taiwan University, No.  
 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan  
 3 (bases 1 to 2592)  
 Hsieh, J.S., Chen, M.R. and Hsing, Y.I.C.  
 Direct Submission  
 Submitted (03-JAN-2001) Agronomy, National Taiwan University, No.  
 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan  
 Sequence update by submitter  
 On Jan 3, 2001 this sequence version replaced gi:5616514.  
 Location/Qualifiers  
 1..2592  
 /organism="Sorghum bicolor"  
 /strain="Kafir 5765-6-1-11-3"  
 /db\_xref="taxon:4558"  
 59..1948  
 /codon\_start=1  
 /product="soluble starch synthase"  
 /protein\_id="RAD45815.2"  
 /db\_xref="GI:12019656"

CDS

alignment\_scores:  
 Quality: 2770.50 Length: 744  
 Ratio: 4.426 Gaps: 15  
 Percent Similarity: 84.140 Percent Identity: 73.925  
 alignment\_block:  
 US-09-674-824-2 x AF168786 ..  
 Align seg 1/1 to: AF168786 from: 1 to: 2592  
 4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20  
 71 TCGCCCGTGGCGCGCGTGCCTC..... 94  
 20 pProAlaThrAlaAlaArgAlaSerAlaCysVal..... 31  
 95 .....GTCCTAGCGGAGCGCGCGGCTTGGGCTTGGGCGCGGCC 137  
 32 .....ValArgAlaArgLeuArg.....ArgLeuAlaArgGly 42  
 138 GGGCGCGGATCGGCGCGCGCGGCTTCCAGCGGCTGGTGGCGCAGG 187  
 43 ArgTyrValAlaGluLeuSerArgGluGlyProAla.....AlaArgPr 57  
 188 CGGTGCGTTCGGAGCTGAGCAGGAGGCTCCCGCGCGCGCGCGCC 237  
 57 oAlaGlnGlnGlnLeuAlaProProLeuValProGlyPheLeuAlaP 74  
 238 GCTGCCACCGCGCTGCTGGCGCGCGCGCTGCTGGCGCGCTTCTCGC 286  
 74 roProProAlaProAlaGlnSerProAlaProThrGlnProProLeu 90  
 287 ..CGCGCGTCCGAGCGCGAGGCTGAGCGGCTGACACCGCGCGCC 334  
 91 ProAspAlaGlyValGlyGluLeuAlaProAspLeuLeuGlyI 107  
 335 CCGGACCGCGCTCGCGGACCTCGGT.....CTCAACCTGAAGGGAT 378  
 107 eAlaGluAspSerIleAspSerIleValAlaAlaSerGluGlnAsp 124  
 379 TGCTGAGGTTCCATCGATGAGACAGTAGTTGTGGCAAGTGAGCAAGAT 428  
 124 erGluIleMetAspAlaAsnGluGlnProGlnAlaLysValThrArgSer 140  
 429 CTGAGATCGTGGTGGAGGAGGAGCTGAGCTAAAGTAACAAAGC 478  
 141 IleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLe 157  
 479 ATTGCTTTTGAATCGCGAAGCTTCTCTTATGCAAGCTTGGGGTCT 528  
 157 uGlyAspValCysGlySerLeuProIleAlaLeuAlaArgGlyHisA 174  
 529 AGGAGATGTTTGGTTCATTGCCAGTTGCTGCTGCTGCTGCTGCTCACC 578  
 174 rgValMetValValMetProArgTyrLeuAsnGlySerSerAspLysAsn 190  
 579 GTGTGATGTTGATGCGCCAGATATTAAATGGTACCTCTGATAAGAT 628  
 191 TyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheG 207  
 629 TACGAAATGCAATTTTACACAAAGACACATTTCGATTTCGATTCTTGG 678  
 207 YGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValAsp 224  
 679 CGGTGAACATGAAGTTACCTTTTCCATGAGTACAGATTTCAGTTGACT 728

DKFRGWFSVPVSHRITAGCDILLMPSPREPCGLNQLYAMOXGTVPVVHATGGLRDT  
 VENFNPEGENGEQGTGWAFAPLTTENMFVDIANCFDIOQAIFLGRAHEGHVKRLH  
 VGPCR".

BASE COUNT 593 a 598 c 705 g 696 t  
 ORIGIN



```
224 rpValpheValAspHisProSerTyrHisArgProGlySerLeuTyrGly 240
|||||
729 GGGTGTGTTGATCATCCCTCATATCACAGACCTGGAAATTTATATGGA 778
|||||
241 AspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCy 257
|||||
779 GATAGTTTGTGCTGTTGGTGAATATGATGTTTGTGATACAGCTCCCTTG 828
|||||
257 sTyxAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyTyrIleT 274
|||||
829 CTATGCTGCATGTGAGGCTCTTTGGTCTTGAATTTGGAGGATATATT 878
|||||
274 yrGlyGlnAsnCysMetPheValValAsnAspTyrPheHisAlaSerLeuVal 290
|||||
879 ATGGACAGAATTGCAATGTTGTTGTTGATGATGGCATGGCCAGCTAGTG 928
|||||
291 ProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSe 307
|||||
929 CCAGTCCTCTCTGCTGCAAAATATAGACCATATGGTGTTTATAAGACTC 978
|||||
307 rArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProA 324
|||||
979 CCGCAGCATTCCTGTAATACATAATTTAGCACATCAGGGTGTAGAGCCTG 1028
|||||
324 laSerThrTyrProAspLeuGlyLeuProGluTyrTyrGlyAlaLeu 340
|||||
1029 CAAGCACATATCCTGACCTTGGTGGTCCACCTCAATGGTATGGAGCTCTG 1078
|||||
341 GluTyrPheProGluTyrPheAlaArgArgHisAlaLeuAspLysGlyCl 357
|||||
1079 GAGTGGTATTCCCTCAATGGGCAAGGAGGCATGCCCTTGCAAGGGTGA 1128
|||||
357 uAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValT 374
|||||
1129 GGCAGTTAATTTTGAAGGTGCGAGTTGTGACAGCAGATCGAATTTGGA 1178
|||||
374 hrValSerGlnGlyTyrSerTyrGluValThrAlaGluGlyGlyGln 390
|||||
1179 CTGTCAAGTAAGGTTATTATCATGGAGGTCCACAACCTGCTCAAGGTGGACAG 1228
|||||
391 GlyLeuAsnGlnLeuLeuSerSerArgLysSerValLeuAsnGlyIleVa 407
|||||
1229 GGTCTCAATGAGCTCTTAAGCTCCCGAAGAGGTGTTAATTAACGGAAATGT 1278
|||||
407 laAsnGlyIleAspIleAsnAspTyrAsnProThrThrAspLysCysLeuP 424
|||||
1279 AAATGGAATGACATTAATGATTTGAACCTCGCAGCGACAAATGATACC 1328
|||||
424 roHisHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysAla 440
|||||
1329 CTTGTCAATATTCTGTGATGACCTCTCGGAAAGCCAAATGTAAAGT 1378
|||||
441 GluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuI 457
|||||
1379 GCATTGCAAGAGAGCTGGGTTTACCTATTAAGGCTGAAGTTCCTCTGAT 1428
|||||
457 eGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysM 474
|||||
1429 TGGTTTATTGGAAGATTGATATCAGAAAGCATTTGATCTCATTTCAAC 1478
|||||
474 etAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuGly 490
|||||
1479 TTATCATACCACATCTCATCGGAGCAGCATCTCAATTTGTGCTTTGGA 1528
|||||
491 SerGlyAspProIlePheGlyTyrPheMetArgSerThrGluSerSery 507
|||||
1529 TCTGGTGACCCAGAGCTCGAAGACTGGATGATCTACAGAGTCGAGCTT 1578
|||||
507 rLysAspLysPheArgGlyTyrPheValGlyPheSerValProValSerHis 524
|||||
1579 CAAGGATAAATTTCTGTGATGGTGGTGGATTTAGTTTCCAGTTTCCACC 1628
|||||
524 rgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluPro 540
|||||
```

```
seq_name: gb_pat:AR106494
seq_documentation_block:
LOCUS AR106494 1752 bp DNA PAT 14-FEB-2001
DEFINITION Sequence 12 from patent US 6107060.
ACCESSION AR106494
VERSION AR106494.1 GI:12821024
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1752)
AUTHORS Keeling, P. and Guan, H.
TITLE Starch encapsulation
JOURNAL Patent: US 6107060-A 12 22-AUG-2000;
FEATURES
location/Qualifiers
1..1752
source
BASE COUNT 433 a 378 c 470 g 471 t
ORIGIN
```



## alignment\_scores:

Quality: 2713.50 Length: 585  
 Ratio: 4.916 Gaps: 3  
 Percent Similarity: 94.359 Percent Identity: 86.325

## alignment\_block:

US-09-674-824-2 x AR106494 ..

Align seg 1/1 to: AR106494 from: 1 to: 1752

45 ValAlaGluLeuSerArgGluGlyProAlaAlaArgProAlaGlnGlnI 61  
 |||||  
 4 GTCCGGAGCTGACAGGGAGGCCCGCCGCGCTGCCACCCGC 53  
 |||||  
 61 nGlnLeuAlaProLeuValProGlyPheLeuAlaProProProA 78  
 |||||  
 54 GCTGCTGGCGCCCGCTCGTGGCGGCTTCCGCG...CCGCGCGCG 100  
 |||||  
 78 laProAlaGlnSerProAlaProThrGlnProProLeuProAspAlaGly 94  
 |||||  
 101 AGCCACGGGTGAGCCGGCATCGACGCCGCGCGCGCTGCCGCGCGC 150  
 |||||  
 95 ValGlyGluLeuAlaProAspLeuLeuLeuGluGlyLeuAlaGluaspSe 111  
 |||||  
 151 CTGGGGACCTCGGT.....CTCGAACCTGAAGGGATTGCTGAAGGTT 194  
 |||||  
 111 rleAspSerIleIleValAlaAlaSerGluGlnAspSerGluIleMeta 128  
 |||||  
 195 CATCGATAACACGACTAGTCTGGCAAGTGAGCAAGATTCTGAGATTGTGG 244  
 |||||  
 128 spLaasnGluGlnProGlnAlaLysValThrArgSerIleValPheVal 144  
 |||||  
 245 TTGGAAAGGAGCAAGCTCGAGCTAAAGTAACACAAAGCATTTGCTTTGTA 294  
 |||||  
 145 ThrGlyGluAlaAlaProTyrAlaLysSerGlyLeuGlyAspValCy 161  
 |||||  
 295 ACCGGCGAAGCTCTCTTAAGCAAGCTGCGGGGCTTAGGAGATGTTG 344  
 |||||  
 161 sGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValy 178  
 |||||  
 345 TGGTTTATGCCAGTTGCTCTTCTGCTGCTGCTGCTACCGGTGATGGTTG 394  
 |||||  
 178 alMetProArgTyrLeuAsnGlySerSerAspLysasnTyrAlaLysAla 194  
 |||||  
 395 TAATGCCAGATATTAAATGATGATACCTCCGATTAAGAAATTATGCAATGCA 444  
 |||||  
 195 LeuTyrThrAlaLysHisIleLysIleProCysPheGlyGlySerHisG 211  
 |||||  
 445 TTTTACACAGAAACACATCTCGATTCCTATGCTTTGGCGGTGAACATGA 494  
 |||||  
 211 uValThrPheHisGluTyrArgAspAsnValAspTyrValPheVala 228  
 |||||  
 495 AGTTACCTCTTCCATGAGTATAGATTCAGTTGACTGGGTGTTGTTG 544  
 |||||  
 228 spHisProSerTyrHisArgProGlySerLeuTyrGlyAspAsnPheGly 244  
 |||||  
 545 ATCATCCCTCATATACACACCTGGAATTTATGAGATGAAGTTGTTG 594  
 |||||  
 245 AlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAlaCy 261  
 |||||  
 595 GCPTTTGGTGATTAATCAGTTTCAGATACACACTCTTGTGATGCTGCATG 644  
 |||||  
 261 sGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGlnAsnC 278  
 |||||  
 645 TGAGGCTCTTTGATCCTTGAATTTGGAGGATATATTTATGGACAGAAAT 694  
 |||||  
 278 ysMetPheValValAsnAspTTPHisAlaSerLeuValProValLeuLeu 294  
 |||||  
 695 GCATGTTTGTTCATGATTGGCATGGCCAGCTAGTGCACAGCTCTCTT 744  
 |||||  
 295 AlaAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLe 311  
 |||||  
 745 GCTGCAAAATATAGACCATATGGTGTTTTAAAGACTCCCGCAGCATCTCT 794

311 uValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrP 328  
 |||||  
 795 TGTAAATACATAAATTTAGCACATCAGGGGTAGAGCCGTGCAAGCACATC 844  
 |||||  
 328 roAspLeuGlyLeuProGluTyrTriPyrGlyAlaLeuGluTyrValPhe 344  
 |||||  
 845 CTGACCTTGGGTGCCACCTGAATGGTATGGAGCTCTGGAGTGGGTATTC 894  
 |||||  
 345 ProGluTyrPalaArgArgHisAlaLeuAspLysGlyGluAlaValAsnPh 361  
 |||||  
 895 CCTGAATGGCGAGGAGGCATGCCCTTGACAAGGGTGAGCAGTAAATTT 944  
 |||||  
 361 eLeuLysGlyAlaValValThrAlaAspArgIleValThrValSerGlnG 378  
 |||||  
 945 TTTGAAGGTGCAGTTGTGACAGCAGATCGAATCGTGACTGTTCAGTAAG 994  
 |||||  
 378 lyTyrSerTyrGluValThrAlaGluGlyGlyGlnGlyLeuAsnGlu 394  
 |||||  
 995 GTTATTCTGGAGGTACAACTGCTCAAGGTGACAGGGCCCTCAATGAG 1044  
 |||||  
 395 LeuLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAs 411  
 |||||  
 1045 CTCTTAAGCTCCAGAAAGAGTGTATTAAACGGAATTTGTAATGGAATGA 1094  
 |||||  
 411 pIleAsnAspTyrAsnProThrThrAspLysCysLeuProHisHisTyrS 428  
 |||||  
 1095 CATTAATGATTGGAAACCTGCCACAGACAAATGTATCCCTGTCTCATATT 1144  
 |||||  
 428 erValAspAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGlnLys 444  
 |||||  
 1145 CTGTTGATGACCTCTCTGGAAGGCCAAATGTAAAGGTGCTATTCGAGAAG 1194  
 |||||  
 445 GluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheIleG 461  
 |||||  
 1195 GAGCTGGGTTCACCTATAAGGCTGATGTCTCTGATTGGCTTTATTGG 1244  
 |||||  
 461 yArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIleProG 478  
 |||||  
 1245 AAGTTGGATTATCAGAAAGGCATTGATCTCACTCAACTTATCATACCG 1294  
 |||||  
 478 luLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAspPro 494  
 |||||  
 1295 ATCTCATGCGGGAAGATGTTCAATTTGTCATGCTGGATCTGGTGACCCA 1344  
 |||||  
 495 IlePheGluGlyTyrMetArgSerThrGluSerSerTyrLysAspLysPh 511  
 |||||  
 1345 GAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCAAGGATAAAT 1394  
 |||||  
 511 eArgGlyTyrPValGlyPheSerValProValSerHisArgIleThrAlaG 528  
 |||||  
 1395 TCGTGGATGGTGGATTAGTTAGTGTCCAGTTTCCACCCACCGAATAACGCG 1444  
 |||||  
 528 lyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLeuAsn 544  
 |||||  
 1445 GCTGCGATATATTGTAATGCCATCCAGATTCGAACCTTGTGTGCTCAAT 1494  
 |||||  
 545 GlnLeuTyrAlaMetClnTyrGlyThrValProValValHisGlyThrG 561  
 |||||  
 1495 CAGCTATATGCTATGAGTATGGCAGTTCCTGTTGCTCCATGCAACTGG 1544  
 |||||  
 561 yGlyLeuArgAspThrValGluThrPheAsnProPheGlyAlaLysGlyG 578  
 |||||  
 1545 GGGCTTAGAGATACCGTGGAGAACTTCAACCTTTCCGGTGAGAAATGGAG 1594  
 |||||  
 578 luGluGlyThrGlyTyrPalaPheSerProLeuThrValAspLysMet.le 594  
 |||||  
 1595 AGCAGGGTACAGGGTGGCATTCGACCCCTAACCCACAGAAACATGTTT 1644  
 |||||  
 594 utrpAlaLeuArgThrAlaMetSerThrPheArgGluHisLysProSerT 611  
 |||||  
 1645 GTGGACATTCGGAAGTGCATATATACATACAGGGAACACAAAGTCTCTCT 1694



611 rpgluGlyLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHis 627  
|||||  
1695 GGAAGGCTAATGAACGAGGATGTCAAAGACTTCACGTGGGACCAT 1744  
628 Ala 628  
|||  
1745 GCC 1747

seq\_name: gb\_pat:A63308

seq\_documentation\_block: 2992 bp DNA PAT 12-MAR-1998  
LOCUS A63308  
DEFINITION Sequence 1 from Patent WO9720936.  
ACCESSION A63308

VERSION A63308.1 GI:3717138

KEYWORDS  
SOURCE unidentified.

ORGANISM unidentified.

unclassified.

REFERENCE 1 (bases 1 to 2992)

AUTHORS Keeling,P.L. and Knight,M.E.

TITLE MODIFICATION OF STARCH SYNTHESIS IN PLANTS

JOURNAL Patent: WO 9720936-A 1 12-JUN-1997;

ZENeca LTD (GB)

COMMENT Other publication AU 1037197 19970627.

FEATURES  
Location/Qualifiers

1..2992

/organism="unidentified"

/db\_xref="taxon:32644"

/clone="NUMBER 1"

BASE COUNT 758 a 655 c 801 g 776 t 2 others

ORIGIN

alignment\_scores:

Quality: 2655.50 Length: 745

Ratio: 4.304 Gaps: 11

Percent Similarity: 82.819 Percent Identity: 71.544

alignment\_block:

US-09-674-824-2 x A63308 ..

Align seg 1/1 to: A63308 from: 1 to: 2992

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20

|||||

614 TCGCGCGTGGCGCGCGCGTGCCTC.....CTCCTCGCGCG 648

20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAlaLeuA 37

|||||

649 GGNGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTCC 698

37 rgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArgGluGlyPro 53

|||||

699 AGCGCGTCTCGCGCGCGCGCGTGCCTCGCGGAGCTGAGCAGGAGGCGCC 748

54 AlaAlaArgProAlaGlnGlnGlnLeuAlaProProLeuValProG1 70

|||||

749 GCGCGCGCGCGCGTCCACCGCGCGTGTGCGCGCGCGCGCGTGTGCGCGG 798

70 yPheLeuAlaProProProProAlaProAlaGlnSerProAlaProThrG 87

|||||

799 CTTCCTCGCG...CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTCC 845

87 InProProLeuProAspAlaGlyValGlyGluLeuAlaProAspLeu 103

|||||

846 CGCGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTCC 889

104 LeuGluGlyLeuAlaGluAspSerIleAspSerIleValAlaAlaSe 120

|||||

890 CCTGAAGGAGTGTGTAAGGTTCATCATCAACACAGTAGTGTGCGAAG 939

120 rGluGlnAspSerGluIleMetAspAlaAsnGluGlnProGlnAlaLysV 137

|||||  
940 TGAGCAAGATTCTCAGATTGTGTGGAAAGGAGCAAGCTCGAGCTAAAG 989  
137 alThrArgSerIleValPheValThrGlyGluAlaAlaPro.TyrAlaLys 153  
|||||  
990 TAACACAAACGATGTCTTGTAAACCGCGAAGCTTCTCCTTAATCGAAA 1039  
153 sSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuAlaA 170  
|||||  
1040 GTCTGGGGTCTAGGAGATGTTGTGGTTCATTGCCAGTTCCTTGTCTG 1089  
170 laargGlyHisArgValMetValValMetProArgTyrLeuAsnGlySer 186  
|||||  
1090 CTCGCGGTACCGGTGTGTTGTAATGCCAGACATTTAAATGGTATACC 1139  
187 SerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLysI 203  
|||||  
1140 TCCGATAAGAAATATGCAATGCATTTTACTCAGAAAACACATTCGGAT 1189  
203 eProCysPheGlyGlySerHisGluValThrPhePheHisGluTyrArg 220  
|||||  
1190 TCCATTCTTTGGCGGTGAACATGAAGTTACCTTCTCCATGAGTATAGAG 1239  
220 spAsnValAspTrpValPheValAspHisProSerTyrHisArgProGly 236  
|||||  
1240 ATTGAGTTGAGTGGTGTGTTGTTGATCATCCTCATATACACACCTGGA 1289  
237 SerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTy 253  
|||||  
1290 AATTATATGAGATAAGTTGGTGTCTTTGGTGATATATCAGTTTCAGATA 1339  
253 rThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuG 270  
|||||  
1340 CACACTCCTTTGCTATGCTGCATGTGAGGCTCCTTTGGTCTTGAATGG 1389  
270 lyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrpHis 286  
|||||  
1390 GAGGATATATTATGGACAGAAATTCATGTTGGTGTTCATGATGGCAT 1439  
287 AlaSerLeuValProValLeuLeuAlaLysTyrArgProTyrGlyIle 303  
|||||  
1440 GCCAGTCTAGAGCCAGTCTCTTCTGCTGCAAAATATAGACCATATGCTGT 1489  
303 lTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisGlnG 320  
|||||  
1490 TTATAAAGACTCCCGCAGCATCTTGTGTAATACATAATTTAGCATCAGG 1539  
320 lyValGluProAlaSerThrTyrProAspLeuGlyLeuProGluTrp 336  
|||||  
1540 GTGTAGAGCTGCAAGCACATATCTGCACCTTGGTGGCCACCTGAATGG 1589  
337 TyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAlaLe 353  
|||||  
1590 TATGGAGCTCTGGAGTGGTATTCCTCAATGGCGGAGGAGGATGCCCT 1639  
353 uAspLysGlyGluAlaValAsnPheLeuLysGlyValAlaValThrAlaA 370  
|||||  
1640 TGACAAGGGTGAGGAGTAAATTTTCAAAAGGTGACAGTGTGTGACAGCAG 1689  
370 spArgIleValThrValSerGlnGlyTyrSerTrpGluValThrThra 386  
|||||  
1690 ATCGAATCGTCACTGTGAGTAAGGTTTATTCATGGAGGTCCACAACGTCT 1739  
387 GluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerValLe 403  
|||||  
1740 GAAGGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCAGAAAGAGTGTATT 1789  
403 uAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrThra 420  
|||||  
1790 AAACGGAATTTGAATGAATTTGACATTAATGTTGAACCTCCACACAG 1839  
420 spLysCysLeuProHisTyrSerValAspAspLeuSerGlyLysAla 436  
|||||



```

1840 ACAATGTATCCCTGTCTCATTATTCTGTTGATGACCTCTCTTGAAAGGCT 1889
437 LysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAs 453
|||||... |||||||... |||||||... |||||||... |||
1890 AATGTAAAGGTGCATTTCGAAAGGAGGTGGGTTTACCTATAAGGCCTGA 1939
453 pValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleA 470
|||||... |||||||... |||||||... |||||||... |||
1940 TGTTCCTCTGATTGGCTTTATTGGAAGATTGGATTATCAGAAAGGCATTG 1989
470 spLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGlnPhe 486
|||||... |||||||... |||||||... |||||||... |||
1990 ATCTCATTCACTATATACAGATCTCATCGGAGATGTTCAATTT 2039
487 ValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgSerTh 503
|||||... |||||||... |||||||... |||||||... |||
2040 GTCATGCTTGGATCTGGTGACCCAGAGCTTGAAGATTGGATGAGATCTAC 2089
503 rGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerValp 520
|||||... |||||||... |||||||... |||||||... |||
2090 AGAGTCGATCTTCAAGGATAAATTCGTGGATGGTTGGATTAGTGTTC 2139
520 roValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSer 536
|||||... |||||||... |||||||... |||||||... |||
2140 CAGTTTCCACCAGATAACTCGGCTGGCGATATTGTTAATGCCATCC 2189
537 ArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyTh 553
|||||... |||||||... |||||||... |||||||... |||
2190 AGATTGGAACCTTGTGCTCAATCAGCTATATGCTATGCAGTATGGCAC 2239
553 rValProValValHisGlyThrGlyGlyLeuArgAspThrValGluThrP 570
|||||... |||||||... |||||||... |||||||... |||
2240 AGTTCCTGTTCTCATGCACTGGGGGCTTAGAGATACCGTGGGCAACT 2289
570 heAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPheSer 586
|||||... |||||||... |||||||... |||||||... |||
2290 TCACCCCTTCGGTCAGANTGGAGAGGAGGTACAGGGTGGGCATTCGCA 2339
587 ProLeuThrValAspLysMet. LeuTrpAlaLeuArgThrAlaMetSert 603
|||||... |||||||... |||||||... |||||||... |||
2340 CCCCTAACCCACAGAAAAATGTTTGTGGACATTGCGAATGCAATATCTA 2389
603 hrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMet 619
|||||... |||||||... |||||||... |||||||... |||
2390 CATACAGGACACACAGTAATAATGGGAGGGCTAATGAAGCCAGGCATG 2439
620 ThrLysAspHisThrTrpAspHisAlaPro. SerSerThrSerArgSerS 636
... |||||||... |||||||... |||||||... |||||||... |||
2440 TCAAAAGAGTTCACGTGGGACCATGGCGCTGAACAATACGACAAATCTT 2489
636 erSerGlyProSerTrpThrAsnProThrSerCysArg. ArgGlyLeuGl 652
|||||... |||||||... |||||||... |||||||... |||
2490 CCAGTGGGCTTCATCGGATCGACCC... GATGTTCAATGGAAAAAGGG 2536
652 yArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSerSerp 669
... ||| ... ||| ... ||| ... ||| ... |||
2537 ACCAAAGTTGGTTGGTTCCTTGAAGATTATCAGTTCATCATCTATAGTA 2586
669 heArgGlyProGluGlyTyPro.....CysThrLeuArgCysProAla 683
... ||| ... ||| ... ||| ... ||| ... |||
2587 AGCTGAATGATGAAGAAAAACCCCTGTACATTACATGGAAAGGCAGACCGG 2636
684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPhe. 695
... ||| ... ||| ... ||| ... ||| ... |||
2637 CTATTGGCTCCATGCTCCCAATGTCGCTTTGGCTTGCCTTCGCTGATGG 2686
695 ..... 695
2687 ACCGGATGCAGTGAGGAATCCAGNCGACGACAGTTTGGAGGATAGGAA 2736
696 .....AlaGlySerArgThrTyrAspGlyCysAlaAlaValThr 710
|||||... |||||||... |||||||... |||||||... |||
2737 GGGAGCTGGAAGC...AGTCACGCGAGGAGGCAAGCCTTCGCGCTTAAT 2783

```

```

711 AlaSerGlyGlyArgGlnLeuGlnPhe 719
... |||
2784 TCATATGGAACAAGCTGGAGTCAGCTTT 2810

```







OM of: US-09-674-824-2 to: N\_Geneseq\_1101:\* out\_format : pfs  
Date: Mar 28, 2002 5:41 PM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:  
-MODEL=frame\_p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO.spool/US09674824/runat\_27032002\_094252\_18454/app\_query.fasta\_1.824  
-DB=N\_Geneseq\_1101 -OFMT=fastap -SURFIX=ring -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09674824 -CGCN1\_1\_261 -NCFU=6 -ICFU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:  
Query: US-09-674-824-2  
Query length: 756  
Database: N\_Geneseq\_1101:\*  
Database sequences: 930621  
Database length: 428662619  
Search time (sec): 180.010000

score_list:	Strd Orig	Zscore	EScore Len	Documentation
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ24487			+ 4044.00 4329.98	4.2e-233
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAZ34651			+ 3930.00 4208.06	2.6e-226
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV01527			+ 3612.00 3868.22	2.2e-207
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50636			+ 2826.50 3023.93	2.3e-160
/SIDSL/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ45183			+ 2821.50 3018.41	2.7e-160
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAZ50785			+ 2809.50 3006.08	2.3e-159
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50643			+ 2806.50 3004.41	2.9e-159
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV29756			+ 2709.50 2901.51	1.5e-153
/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT:AAV70960			+ 2708.50 2900.46	1.8e-153
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV67285			+ 2655.50 2838.67	9.0e-150
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV68832			+ 2651.00 2833.85	9.0e-150
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAV29759			+ 2346.50 2727.23	7.9e-144
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50637			+ 2495.00 2672.47	8.8e-141
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ34652			+ 2349.00 2498.33	4.4e-131
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50642			+ 2036.00 2180.38	2.3e-113
/SIDSL/gcgdata/geneseq/geneseq/NA1996.DAT:AAZ46797			+ 1854.00 1981.90	2.6e-102
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV32325			+ 1731.50 1851.49	4.7e-95
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV70958			+ 1252.50 1334.31	3.0e-66
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50651			+ 1246.00 1328.64	6.3e-66
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50646			+ 1239.50 1321.03	1.7e-65
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1212.00 1292.48	8.5e-64
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1209.00 1290.32	6.5e-64
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1203.00 1279.41	3.5e-63
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1203.00 1279.41	3.5e-63
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1200.50 1277.14	4.6e-63
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1200.50 1277.14	4.6e-63
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1182.50 1257.75	5.6e-62
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1154.50 1230.36	1.9e-60
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1154.50 1230.36	1.9e-60
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50647			+ 1076.00 1146.53	8.7e-56
/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT:AAV29754			+ 1076.00 1146.53	8.7e-56
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV67286			+ 1071.00 1140.81	1.8e-55
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV66833			+ 1071.00 1140.81	1.8e-55
/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT:AAV67287			+ 943.50 1002.72	8.9e-48
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV29753			+ 943.50 1002.72	8.9e-48
/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT:AAV67287			+ 900.00 955.42	3.9e-45
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV67287			+ 887.50 943.04	1.9e-44
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV52178			+ 881.00 921.39	3.0e-43
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ506319			+ 870.50 926.32	1.6e-43
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ506319			+ 863.00 917.37	5.1e-43

/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAZ36227	-	841.50	892.51	1.2e-41
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ36227	+	788.00	823.65	8.4e-38
/SIDSL/gcgdata/geneseq/geneseq/NA1992.DAT:AAQ25978	+	678.50	722.63	3.6e-32
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV23619	+	678.50	722.63	3.6e-32
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV70951	+	634.50	675.24	1.6e-29
seq_name: /SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ24487				
seq_documentation_block:				
ID	AAZ24487 standard; DNA; 2805 BP.			
XX				
AC	AAZ24487;			
XX				
DT	18-FEB-2000 (first entry)			
XX				
DE	Wheat soluble starch synthase DNA.			
XX				
KW	Soluble; starch synthase; wheat; transgenic plant; starch production;			
KW	food; baking; pastry; packaging material; glucose; glucan; paper; pulp;			
KW	adhesive; textile; building material; soil stabilizer; wetting agent;			
KW	fertilizer; plant-protection; cosmetic; flocculant; ss.			
XX				
OS	Triticum aestivum.			
XX				
FH	Key	Location/Qualifiers		
FT	CDS	314..2584		
FT		/*tag= a		
FT		/product= "soluble starch synthase"		
XX				
PN	DE19820607-A1.			
XX				
PD	11-NOV-1999.			
XX				
PF	08-MAY-1998; 98DE-1020607.			
XX				
PR	08-MAY-1998; 98DE-1020607.			
XX				
PA	(AGRE ) HOECHST-SCHERING AGREVO GMBH.			
XX				
PI	Loerz H, Lueticke S, Block M;			
DR	WPI; 2000-024508/03.			
XX	P-PSDB; AAY50818.			
PT	New enzyme with starch synthase activity, useful for producing starch			
XX	for foods and packaging materials -			
XX				
PS	Claim 1b; Page 15-19; 24pp; German.			
XX				
CC	This invention describes a novel protein (I) with the activity of wheat			
CC	starch synthase. Transgenic plants, specifically wheat, that contain (I)			
CC	are used for production of starch, used particularly in foods,			
CC	particularly baked and pastry goods and for making packaging materials or			
CC	disposable items. Starch may also be used as starting materials for			
CC	glucose or glucan components (e.g. for fermentation or further chemical			
CC	conversion); in paper and pulp production, as adhesives, in textiles,			
CC	in preparation of gypsum-based building materials, as soil stabilizer,			
CC	as wetting agent etc. in fertilizer and plant-protection compositions,			
CC	as binder (in pharmaceuticals, cosmetics, coal briquetting and casting			
CC	sand), as flocculant in soil or coal slurries, as rubber and leather			
CC	additives, and for production of synthetic polymers, e.g. polyurethane			
CC	films. Transgenic plants with increased/decreased production of (I)			
CC	produce starches with altered physical and/or chemical properties such as			
CC	amylose/amylopectin ratios, degree of branching, mean chain length,			
CC	phosphate content, gelatinization properties, gel- or film-forming			
CC	properties, or starch grain size or structure. This sequence encodes the			
CC	soluble starch synthase isolated from wheat (Triticum aestivum L. cv.			
XX	Florida).			
SQ	Sequence '2805 BP; 683 A; 703 C; 763 G; 656 T; 0 other;			
	alignment_scores:			



Quality: 4044.00 Length: 756  
 Ratio: 5.349 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-674-824-2 x AA24487 ..

Align seg 1/1 to: AA24487 from: 1 to: 2805

1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17  
 314 ATGGCGGCGAGCGGCGCTCGGCGCGGCGGCTCGCCCGCCAGCGCTCCGCT 363  
 17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34  
 364 GCGCGCGGATCCGCGGACGCGGCGGCGGCGGCTCGGCTCGCTCGCGG 413  
 34 laArgLeuArgArgAlaAlaArgGlyArgTyrValAlaGluLeuSerArg 50  
 414 CGCGGCTCCGCGCTTGGCGGCGGCGCTACGTCGCGGAGCTCAGCAGG 463  
 51 GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 67  
 464 GAGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 513  
 67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84  
 514 CGTGCCAGGCTTCCTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 563  
 84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100  
 564 CCCGACGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 613  
 101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117  
 614 GACCTCCCTGCTCGAAGGAGTTCGTGAGATTCCATCGACACATAATGT 663  
 117 AlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134  
 664 GCGTCAAGTAGCAGGAGTTCGTGAGATCATGATCGGAATAGCAACCTC 713  
 134 InAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150  
 714 AAGCTAAAGTTACAGTACGTCGTTGTTGACGTTGTTGACGTTGTTGAC 763  
 151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167  
 764 TATGCAAACTCAGGGGGTTGGGAGATGTTGTTGCTTACCAATTCG 813  
 167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184  
 814 TCTTGCTGCTGCTGCTACCGAGTATGTTGTTGTTGTTGTTGTTGTTGTT 863  
 184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200  
 864 ATGGGTCCTCTCATATAAACTATGCAAGGCATATACACTCGGAAGCAC 913  
 201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisGl 217  
 914 ATTAAGATTCCATGCTTTGGGGGATCAGTGAAGTGAACCTTTTTCATCA 963  
 217 uTyrArgAspAsnValAspTyrValPheValAlaAspHisProSerTyrHis 234  
 964 GTATAGAGACAAGCTCGATTGGGTGTTTGTGATCATCCGTCATATCACA 1013  
 234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250  
 1014 GACCAGGAAGTTTATATGAGATAAATTTGGTGGCTTTTGGTATATATCAG 1063  
 251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267  
 1064 TTCAGATACACACTCCTTTGCTATGCTGATGCGAGGCGGCGGCGGCGGCT 1113

267 uGluLeuGlyGlyTyrIleTyrGlyClnAsnCysMetPheValValAsnA 284  
 1114 TGAATTGGGAGGATATATTTATGGCAGAAATTCATGTTGTTGTTGTAAGC 1163  
 284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300  
 1164 ATTGGCATGCCAGCCTGTGCCAGTCTCTTCTGCTGCAAAATATAGACCA 1213  
 301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317  
 1214 TAGCGTGTTCACAGATTCGCCGACGCCCTTGTATACATAAATTAGC 1263  
 317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334  
 1264 ACATCAGGGTGTGGAGCCTGCAAGTACATATCTGATCTGGGATTCGCTC 1313  
 334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350  
 1314 CTGAATGGTATGAGCTTTAGAAATGGTATTTCCACAATGGGCAAGGAG 1363  
 351 HisAlaLeuAspLysGlyGluAlaValAlaAsnPheLeuLysGlyAlaValVa 367  
 1364 CATGCCCTTCACAAGGTGAGGAGTTAACTTTTGAAGGAGCAGCTGT 1413  
 367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384  
 1414 GACAGCAGATCGGATGTGACCGTCAGTCAGGTTATTCATGGGAGGTCA 1463  
 384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuSerSerArgLys 400  
 1464 CAATGCTGAAGGTGGACAGGCTCAATCAGCTCTTAAGCTCCGCAAAA 1513  
 401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417  
 1514 AGTGTATTGAATGGAATGTAAATGGAATTGACATTAATGATTGGAAACC 1563  
 417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434  
 1564 CACCACAGACAAAGTCTCTCTCTCATCATATTCTGTCGATGACCTCTCTG 1613  
 434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450  
 1614 GAAAGGCCAAATGTAAAGCTGAATTCAGAAAGGAGTTGGGTTTACCTGTA 1663  
 451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467  
 1664 AGGAGGAGATGTTCTCTGATTGGCTTTTATGGAAGACTGGATTACCAAGAA 1713  
 467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484  
 1714 AGGCATTGATCTCATTTAAATGGCCATTCCAGAGCTCATAGGAGGAGGACG 1763  
 484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500  
 1764 TGCAATTTGTCATGCTTGGATCTGGGATCCCAATTTTGAAGGCTGGATG 1813  
 501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517  
 1814 AGATCTACCGAGTCGAGTTACAAGGATAAATTCGCTGGATGGGTGGATT 1863  
 517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534  
 1864 TAGCTTCCAGTTTCCACAGAAATAACTGCAGGTTGCGATATATGTTTAA 1913  
 534 eProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550  
 1914 TGCCATCGAGATTGCAACCTTCGCGCTCTTAATCAGCTATATGCTATGCAA 1963  
 551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567  
 1964 TATGTTACAGTTCTCTAGTTTCATGGAACCTGGGGGCTCCGAGACACAGT 2013  
 567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584



```
|||||
2014 CGAGACCTTCAACCTTTTGTGCAAAAGGAGGAGGTACAGGTGG 2063
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
2064 CGTTCTCACCGCTAACCGTGGACAAGATGTTGTGGCATTCGGAACCGG 2113
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
2114 ATGTCGACATTGAGGAGCAAGCCGTCCTGGAGGGGCTCATGAAGCG 2163
617 gGlyMetThrLysAspHisThrTrpAspHisAlaProSerSerThrSera 634
2164 AGGCATGACGAAGAACCATACGTGGGACCATCCCGGAGCAGTACGAGCA 2213
634 rgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgGly 650
2214 GATCTTCAGTGGGCTTCGTGGACCAACCCCTACGTATGATGAGCGGGA 2263
651 LeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSe 667
2264 CTGGGAGGTCCAAGTGGAGTCTCTTCACGCTCTGAAGACATCTCTTC 2313
667 rSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAlar 684
2314 ATCTCTCCGCGCGCCGGAAGGATACCCCTGTACATTCGTTGTCTGCTA 2363
684 hrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThr 700
2364 CAGTAGAGTCGCANTGCGCTGCTGTGTTGGTTCCGCGGTTCGAGAACA 2413
701 TyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnLe 717
2414 TATGACGGCTGTGCTGTGCGCGGTGACACGCTTCGGGTGACGACAGTT 2463
717 uGlnPheTrpGlyTleArgLysGlyCysAlaAlaGlyTrpLeuThrAlaL 734
2464 ACAGTTTGGGAATAAGGAGGATGCTGTGAGGATGTTTAAACAGCAA 2513
734 ysHisHisSerAspGlySerLeuSerValArgValThrAlaGluIleArg 750
2514 AGCACCACCTCAGATGGCAGCTCTCTGTGCGTGTACAGCTGAANTCAGA 2563
751 AsnGlnLeuValThrLeu 756
2564 AACCAACTGGTGACTCTT 2581
```

seq\_name: /STD51/gcgdata/geneseq/geneseqn/NA1999.DAT:AA34651

seq\_documentation\_block:

ID AAX34651 standard; cDNA; 2662 BP.

XX AC AAX34651;

XX DT 05-JUL-1999 (first entry)

XX DE cDNA sequence of wheat starch soluble synthase I (SSS I).

XX KW Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBSS;  
KW starch branching enzyme; starch soluble synthase; debranching enzyme;  
KW endosperm; wheat; barley; granule-bound synthase; glutenin; starch;  
KW grain softness protein I; bacterial isoamylase; glycogen synthase;  
KW WSBE I-D4 gene; ss.

XX OS Triticum tauschii.

XX PN WO9914314-Al.

XX PD 25-MAR-1999.

XX PF 11-SEP-1998; 98WO-AU00743.

XX PR 20-MAR-1998; 98AU-0002509.

PR 12-SEP-1997; 97AU-0009108.  
XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (GOOD-) GOODMAN FIELDER LTD  
PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.  
PA (AUSU ) UNIV AUSTRALIAN NAT.

XX Li Z, Morell M, Rahman S;  
XX WPI; 1999-229525/19.

XX New isolated cereal plant enzyme genes used for, e.g. expression of  
XX antisense sequences of granule bound synthase

PS Claim 10; Page 81-83; 171pp; English.

XX The invention relates to a novel enzyme of starch biosynthetic pathway  
CC in a cereal plant, where the enzyme is selected from starch branching  
CC enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching  
CC enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or  
CC SBE I of rice or maize. The methods and products can be used for  
CC targeting expression specifically to the endosperm of the seeds of cereal  
CC plants such as wheat or barley. They can be used for the expression of  
CC e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low  
CC mol. wt. glutenin, grain softness protein I, bacterial isoamylase,  
CC bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They  
CC can be used for modifying the characteristics of starch produced by a  
CC plant. The present sequence represents the wheat SSS I cDNA sequence.

XX Sequence 2662 BP; 592 A; 677 C; 754 G; 639 T; 0 other;

alignment\_scores:

Quality: 3930.00 Length: 757

Ratio: 5.247 Gaps: 1

Percent Similarity: 98.943 Percent Identity: 98.415

alignment\_block:

US-09-674-824-2 x AAX34651 ..

Align seg 1/1 to: AAX34651 from: 1 to: 2662

1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17  
|||||  
257 ATGGCGCGGACGCGGCTCGCGCGCGGTCCTCGCCCCCAGCGTCGCGCT 306

17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34  
|||||  
307 GCGCGCGGATCGCGGACGCGCGCGCGCGCTCGCGCTCGCTCGCGCGG 356

34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50  
|||||  
357 CGCGGCTCGCGGCTTGGCGCGCGCGCGCTACGTTGCGGAGCTCAGCAGG 406

51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProle 67  
|||||  
407 GAGGCGCGCGCGCGCGCGCGCGCGCGAGCAGCACTGGCCCCCGCGCT 456

67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84  
|||||  
457 CGTGCGAGGCTTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 506

84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100  
|||||  
507 CCCCCAGCAGCGCGCGCTGCGGACGCGCGGTGGGGAACTCGCGGCC 556

101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117  
|||||  
557 GACCTCTGCTCGAAGGAGTTCCTGAGGATTCATCGACACGACATAATTGT 606

117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProg 134  
|||||  
607 GGTCGAAGTGGCAGGATTCATGATCATGATGATGATGATGATGATGATGAT 656



134 lnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150  
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167  
167 TATGCAAACTCAGGGGGCTGGGAGATGTTGTGGTTCTTACCAATTGC 756  
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184  
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200  
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200  
807 ATGGGTCTCTGATGAAACATATGCAAGGCAATATATACCTGCGAGCAC 856  
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG 217  
857 ATTAAGATTCCATGCTTTGGGGATCACATGAAGTGACCTTTTTCATGA 906  
217 uTyrArgAspAsnValAspTyrValPheValAspHisProSerTyrHisA 234  
907 GTATAGACAAACGTCGATGGGTGTTGTCTGATCATCCGTCATATCATA 956  
234 tGProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250  
957 GACCAAGGAATTTATATGAGATATTTTGGTGTCTTTGGTGATAATCAG 1006  
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267  
1007 TTCAGATACACACTCTCTTTGCTATGCTGCATCGAGAGCCCACTAATCT 1056  
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284  
1057 TGAATTGGGAGGATATTTATGGACAGAAATGTCATGTTGTGTGAACG 1106  
284 sPTPrHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300  
1107 ATTGGCATCCAGCCTTGTGCGAGTCTCTTCTGCTGCAAAATATAGACCA 1156  
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAlaLeuAl 317  
1157 TACGGTGTTTACAGAGATTTCCCGCAGCACCTTGTATACATTAATTTAGC 1206  
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProP 334  
1207 ACATCAGGCTCTGGAGCCTGCAAGTACATATCTGATCTGGGATTGCCAC 1256  
334 roGluTyrTyrGlyAlaLeuGluTyrValPheProGluTyrAlaArgArg 350  
1257 CTGAATGGTATGGAGCTTTAGAATGGGTATTTCCAGAAATGGCAAGGAGG 1306  
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367  
1307 CATGCCCTTGCAGAGGCTGAGCGAGTTAACTTTTGAAGGAGCAGCTCGT 1356  
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTyrGluValT 384  
1357 GACAGCAGATCGAATTTGTGACCGCTCAGTCAGGCTTATTCATGGAGGTCA 1406  
384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400  
1407 CAACCTGCTGAAGGTGACAGGGCTCAATGAGCTCTTAAGCTCCCGGAAA 1456  
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTyrAsnPr 417  
1457 AGTGTATTTGAATGGAATTTGAATGGAATTTGACATTAATGATTTGGAACCC 1506  
417 oThrThrAspLysCysLeuProHisIstYrSerValAspAspLeuSerG 434  
1507 CACCACAGACAAGTGTCTCCCTCATCATTTCTGTGATGACCTCTCTG 1556

434 lYlYsAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450  
1557 GAAAGCCAAATCTAAAGCTGAATTCAGAAAGGAGCTGGGTTTACCTGTA 1606  
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467  
1607 AGGAGGAGATGTTCTCTGATTGGCTTTATTTGGAAGACTGGATTACCAGAA 1656  
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484  
1657 AGCATTTGATCTCATTAANATGCCATTCCAGAGCTCATGAGGAGGAGCG 1706  
484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTyrMet 500  
1707 TGCAGTTTGTCTGATCTGGATCTGGGATCCAAATTTTGAAGCTGGATG 1756  
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTyrValGlyPh 517  
1757 AGATCTTACCGAGTCGAGTTACAGGATAAATTCCTGGATGGGTGGATT 1806  
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534  
1807 TAGTGTTCAGATTTCACACAGAATAACTGCAGGTTGCGATATATTGTAA 1856  
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550  
1857 TGCCATCCAGGTTTGAACCTTGTGCTTAAATCAGCTATATGCTATGCAA 1906  
551 TyrGlyThrValProValHisGlyThrGlyGlyLeuArgAspThrVa 567  
1907 TATGGTACAGTTCCTGTAGTTTCATGGAACCTGGGGCTCCGAGACACAGT 1956  
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584  
1957 CGAGACCTTCAACCTTTTGTGTCAAAAGGAGAGAGGGTACAGGGTGGG 2006  
584 lApeSerProLeuThrValAspLysMetLeuTyrAlaLeuArgThrAla 600  
2007 CGTTCCTACCCCTTACCGTGGACAAAGATGTTGTGGCATTCGGAACCGCG 2056  
601 MetSerThrPheArgGluHisLysProSerTyrGluGlyLeuMetLysAr 617  
2057 ATGTCGACATTCAGGGAGCACAAAGCCGCTCTGGAGGGGCTCATCAAGCG 2106  
617 gGlyMetThrLysAspHisThrTrpAspHisAlaProSerSerThrSer 633  
2107 AGGCATGACGAAAGACCATACGTGGACCATGCCCCCGAGCAGTACGAGC 2156  
634 ArgSerSerSerGlyProSerTyrTrpThrAsnProThrSerCysArgG 650  
2157 AGATCTTCAATGGGCTTCTGTGGACCAACCTACGTCATGACGAGGGG 2206  
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667  
2207 ACTGGGAGGTGCAAGCGCGGCTCTCTTGGAGCTCTGAAGACATGTTCT 2256  
667 erSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAla 683  
2257 CATCTTCCGGCGCCCGAAGGATACCCCTGTACATTGCGTTGCTCTCT 2306  
684 ThrValGluSerGlnCysAlaCysLeuLeuTyrPheAlaGlySerArgTh 700  
2307 ACAGTAGAGTCGCAATGCGCTGCTTGTCT.TGGTCCCGCGGTTCGAGAGT 2355  
700 rTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnL 717  
2356 ACATCAGGCTGTGCTGCTGCGCGGCTGACAGCTTCGGGTGGATGACACT 2405  
717 euGlnPheTyrGlyIleArgLysGlyCysAlaAlaGlyTyrLeuThrAla 733  
2406 TACAGTTTGGGGAATAGCAAGGATGTGCTGACAGGATGGTTTACAGCA 2455  
734 LysHisHisSerAspGlySerLeuSerValArgValThrAlaGluIleAr 750



|||||  
2456 AAGCACCACTCAGATGGCAGCCTCTCTCGTGTATACAGTGAATCAG 2505  
750 gAsnGlnLeuValThrLeu 756  
|||||  
2506 AAACCAACTGGTACTCTT 2524

seq\_name: /STD1/gcgdata/geneseq/geneseq/NA1998.DAT:AAV01527

seq\_documentation\_block:

ID AAV01527 standard; cDNA to mRNA; 2239 BP.

XX AAV01527;

XX 21-MAY-1998 (first entry)

XX Wheat soluble starch synthase partial cDNA sequence.

XX Starch synthase; wheat; transgenic plant; ss.

XX Triticum aestivum L. cv. Florida.

XX Key Location/Qualifiers

FT CDS 3-2018

FT /\*tag= a

PN W09745545-Al.

XX 04-DEC-1997.

XX 28-MAY-1997; 97WO-EP02793.

XX 11-SEP-1996; 96DE-1036917.

XX 29-MAY-1996; 96DE-1021588.

XX (AGRE ) HOECHST-SCHERING AGREVO GMBH.

XX Block M, Loerz H, Luetticke S, Froberg C, Kossmann J;

XX Walter L;

XX WPI; 1998-032652/03.

XX P-PSDB; AAW23937.

PT Nucleic acid encoding starch synthase enzymes from wheat - for  
PT transgenic plants that produce modified forms of starch, useful e.g.  
PT in foods, or for production of packaging materials and disposable  
PT goods

PS Claim 1; Page 47-51; 7lpp; English.

CC This near full-length cDNA clone, designated TASSS, codes for a  
CC soluble starch synthase (see AAW23837) of summer wheat (cv. Florida).  
CC It was isolated from a phage cDNA library of 21-day-old wheat  
CC caryopses by screening with a PCR fragment derived from rice soluble  
CC starch synthase (see also AAV01529-30). A second clone (see AAV01528),  
CC coding for wheat granule-bound starch synthase (see AAW23938) is also  
CC claimed. These isolated nucleic acids can be inserted into vectors  
CC for production of transgenic plants, particularly starch-producing  
CC plants, specifically wheat. Use of the isolated nucleic acids, or  
CC of antisense sequences, allows starch metabolism to be regulated in  
CC transgenic plants. Overexpression may result in improved crop  
CC yield, while modification of starch in planta may eliminate the  
CC need for subsequent chemical/physical modification. Plants with  
CC altered levels of the various isoforms of starch synthase will  
CC produce starch of different chain length, amylose/amylopectin ratio,  
CC degree of branching, phosphate content, gelatinisation behaviour,  
CC granule size and shape, viscosity etc. The starch produced by such  
CC plants is useful particularly in foods or to produce packaging  
CC materials or disposable goods, as well as in any other known use of  
CC starch.

XX Sequence 2239 BP; 611 A; 448 C; 590 G; 590 T; 0 other;

alignment\_scores:

Quality: 3612.00 Length: 671  
Ratio: 5.383 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-674-824-2 x AAV01527 ..

Align seg 1/1 to: AAV01527 from: 1 to: 2239

86 ThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaProAspLe 102  
|||||  
3 ACGCAGCCGCCCTCGCGGACCGCGCTGGGGAACCTCGCGCCGACCT 52  
|||||  
102 uLeuLeuGluGlyLeuAlaGluAspSerIleAspSerIleValAla 119  
|||||  
53 CCTGCTCGAAGGGATTGCTGAGGATCCATCGACAGCATAAATTGGCTG 102  
|||||  
119 laSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProGlnAla 135  
|||||  
103 CAAGTGAGCAGGATTCTGAGATCATGATGCGGAATGAGCAACCTCAAGCT 152  
|||||  
136 LysValThrArgSerIleValPheValThrGlyGluAlaAlaProTyrAl 152  
|||||  
153 AAAGTTACACGTAGCATCGTGTGTGACTGGTGAAGCTGCTCCTTATGC 202  
|||||  
152 aLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeu 169  
|||||  
203 AAAGTCAGGGGGTTGGGAGATGTTGTGTTGCTTACCAATTGCTCTTG 252  
|||||  
169 laAlaArgGlyHisArgValMetValValMetProArgTyrLeuAsnGly 185  
|||||  
253 CTGCTCGTGTACCCAGAGTGTGTTGTAATGCCAAGATACCTTAATGGG 302  
|||||  
186 SerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLy 202  
|||||  
303 TCCTCTGATAAAACATATGCAAGGCAATATACACTCGGAAGCACATTAA 352  
|||||  
202 sileProCysPheGlyGlySerHisGluValThrPhePheHisGluTyr 219  
|||||  
353 GATTCCATGCTTTGGGGGATCACATGAAGTGACCTTTTTCATGAGTATA 402  
|||||  
219 rgAspAsnValAspTrpValPheValAspHisProSerTyrHisArgPro 235  
|||||  
403 GAGACAACGTCGATTGGGTGTTGTCATCATCCGTCATATCACAGACCA 452  
|||||  
236 GlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheAr 252  
|||||  
453 GGAAGTTTATATGGAGATAATTTTGGTGTCTTTTGGTGATAATCAGTTCAG 502  
|||||  
252 gTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluL 269  
|||||  
503 ATACACACTCCTTTGCTATGCTCATCGAGGCCGCCCACTAATCCTGAAT 552  
|||||  
269 euGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrp 285  
|||||  
553 TGGGAGGATATATTTATGGACAGAAATTCATGTTTGTGTGAACGATTGG 602  
|||||  
286 HisAlaSerLeuValProValLeuLeuAlaLysTyrArgProTyrG1 302  
|||||  
603 CATGCCAGCCTTGTGCCAGTCTCTTGTGTCAAAATATAGACCATACGG 652  
|||||  
302 yValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisG 319  
|||||  
653 TGTTTACAGAGATTCGCCGAGCACCTCTGTATACATAATTTAGCACATC 702  
|||||  
319 InGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProGlu 335  
|||||  
703 AGGGTGGAGCCTGCAAGTACATATCTCATCTGGGATTGCTCTCTGAA 752  
|||||  
336 TrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgHisAl 352  
|||||



753 TGGTATGAGCTTTAGAAATGGGTATTTCCAGAAATGGGCAAGGAGGCATGC 802  
352 aLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValValThrA 369  
803 CCTTGACAAGGGTCAGCAGCTAACTTTTGAAGAGGAGCAGTTGTGACAG 852  
369 laAspArgIleValThrValSerGlnGlyTyrSerTrpGluValThrThr 385  
853 CAGATCGGATTGTGACCGTCAGTCAGGGTTATTTCATGGAGGTCACAACT 902  
386 AlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerVa 402  
903 GCTGAAGGTGACAGGGCCTCAATGAGCTCTTAAGCTCCCGAAAAGTGT 952  
402 lLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThr 419  
953 ATTGAATGGAATTCTAAATGCAATTGACATTAATGATTGGAACCCACCA 1002  
419 hrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerGlyLys 435  
1003 CAGACAAGTGTCTCCCTCATCATTAATTTCTGTCGATGACCTCTCTGGAAG 1052  
436 AlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgG1 452  
1053 CCCAAATGTAAGCTGAATCCAGAAGAGTTGGTTTACCTGTAAAGGA 1102  
452 uAspValProLeuLeuGlyPheIleGlyArgLeuAspTyrGlnLysGlyI 469  
1103 GGATGTTCTCTGATTTGGCTTTATTGGAAGACTGGATTACCAGAAGGCA 1152  
469 leAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGln 485  
1153 TTGATCTCATTAATAATGCCATTCAGAGCTCATGAGGAGGAGCGTGCAA 1202  
486 PheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgse 502  
1203 TTTGTCATGCTTGGATCTGGGGATCCAAATTTTGNAGGCTGGATGAGATC 1252  
502 rThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerV 519  
1253 TACCGAGTCGAGTTACAAAGATAAATTCGTTGGATGGGTTGGATTTAGTG 1302  
519 alProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetPro 535  
1303 TTCCAGTTTCCACAGAAATACTGAGGTTGGGATATATTGTTAATGCCA 1352  
536 SerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrG1 552  
1353 TCGAGATTGAACCTTGGGCTTTAATCAGCTATATGCTATGCAATATGG 1402  
552 yThrValProValValHisGlyThrGlyGlyLeuArgAspThrValGluT 569  
1403 TACAGTTCTGTAGTTTCAATGGAACCTGGGGGCTCCGAGACACAGTCGAGA 1452  
569 hrPreAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPhe 585  
1453 CCTTCAACCCCTTTGGTGCACAAAGAGAGAGGGGTACAGGGTGGGGCTTC 1502  
586 SerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSe 602  
1503 TCACCGCTTAACCGTGACAAAGATGTTGTGGCATTCGGAACCGCATGTC 1552  
602 rThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyM 619  
1553 GACATTCAGGGAGCAACAGCGCTCTGGAGGGGCTCATGAAGCGAGGCA 1602  
619 etThrLysAspHisThrTrpAspHisAlaProSerSerThrSerArgSer 635  
1603 TCACGAAAGACCATACGTGGGACCATGCCCGAGCAGTACGAGCAGATCT 1652  
636 SerSerGlyProSerTrpThrAsnProThrSerCysArgArgGlyLeuG1 652  
1653 TCGAGTGGGCTCTCTGGACCAACCCCTACGTCATGTAGACGGGACTGGG 1702

652 yArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSerSerP 669  
1703 GAGTCCCAAGTGCAGTCTCTCTCAGCTCTGAAGACATCTCTTCATCCT 1752  
669 heArgGlyProGluGlyTyrProCysThrLeuArgCysProAlaThrVal 685  
1753 TCCGCGGCGCCGGAAGGATACCCCTGTACATTCGCTTCTCTCTACAGTA 1802  
686 GluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrTyrAs 702  
1803 GAGTCGCAATGCGCTCTGCTTTGTTGTCGCGGTTTCGAGAACATATGA 1852  
702 pGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnLeuGlnP 719  
1853 CGGCTGTGCTGCTGCGCGGTGACAGCTTCGGGTGGACAGATTACAGT 1902  
719 heTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAlaLysHis 735  
1903 TTTGGGAATTAAGAGGAGGATGCTGTCAGAGTGGTTAACAGCAAGCAC 1952  
736 HisSerAspGlySerLeuSerValArgValThrAlaGluIleArgAsnG1 752  
1953 CACTCAGATGCGAGCCTCTCTGTCGCTTACAGCTGAAATCAGAAACCA 2002  
752 nLeuValThrLeu 756  
2003 ACTGGTGACTCTT 2015

seq\_name: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT:AA250636

seq\_documentation\_block:

ID AA250636 standard; cDNA; 2491 BP.

XX AA250636;

XX 23-MAY-2000 (first entry)

XX Corn soluble starch synthase composite gene sequence.

DE Soluble starch synthase; starch fine structure; corn;

XX transgenic plant; amylose; amylopectin; amylose polymerisation;

KW non-granule bound starch synthase; non-GBSSI; altered starch; food;

KW paper; plastic; adhesive; ss.

XX Zea mays.

OS WO200006755-A2.

XX 10-FEB-2000.

XX 26-JUL-1999; 99WO-US16296.

XX 28-JUL-1998; 98US-0094436.

XX (DUPO ) DU PONT DE NEMOURS & CO E. I.

XX Broglie KE, Lightner JE;

XX WPI; 2000-195311/17.

XX Producing transgenic cereal crops with altered starch structure useful

PT for preparing foodstuff, paper, plastic or adhesives, comprises

PT transforming crops with chimeric sense or antisense gene construct

PT encoding starch synthase

XX Claim 5; Page 51; 56pp; English.

XX The present sequence is the corn soluble starch synthase (SSI) composite

CC gene. This was used in the construction of plasmid pSSI1 for the

CC generation of an antisense construct for suppression of SSI expression

CC in corn. The starch fine structure derived from a grain of the cereal

CC crop can be altered in the transformed cereal crop by changes in amylose







```
1674 TAGTGTTCAGTTTCCACCCGAATAAATCGCGCTGGCATATATTGTAA 1723
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
1724 TGCCATCCAGATTCCAACCTTGTGCTCAATACGCTATATGCTATGCAG 1773
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567
1774 TATGCCACAGTTCCTGTGTCCATCAACTGGGGCCTTAGAGATACCGT 1823
567 lgluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrTrpA 584
1824 GGAGAACTCAACCTTTCCGTGAGATGGAGAGCAGCGGTACAGGTGG 1873
584 laphSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
1874 CATTCGCACCCCTTAACCAAGAAACATGTTGTGGACATTCCGAACGTGA 1923
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
1924 ATATCTACATACAGGAACACAAAGTCTCTCTGGGAAGGCTAATGAAGCG 1973
617 gGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSer 633
1974 AGGCATGTCAAAAGACTTCACGTGGGACCATGCCGCTGAACAATACGAAC 2023
634 ArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgL 650
2024 AAATCTTCAGTGGGCCCTTCATCGATCAGCCCTATGTCATGTGAANAAA.. 2071
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSers 667
2072 .....AGGACCAAGTGGTGGTTCCT.....TGAAGAT 2099
667 erSerPheArgGlyPro.....GluGlyTyrProCys 677
2100 CATCATGTCATCATCTCTAGTAACTAATGATGAAAGAAACCCCTGT 2149
678 ThrLeuArgCysProAlaThrValGluSerGlnCysAlaCysLeuTr 694
2150 ACATTACAT..... 2158
694 pPheAlaGlySerArgThrTyrAspGlyCysAlaAlaAlaValThrA 711
2158 ..... 2158
711 laseRgLyGlyArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAla 727
2159 .....GGAAGGAGACCGCGCTATTGGCTCCATTGCTCCAACGCTCTGC 2200
728 AlaGlyTrpLeuThrAlaLysHis 735
2201 TTTGGCTGGCTTGCCTCGATGCAC 2224
seq_name: /SIDSI/gcgdata/geneseq/geneseq/NAL194.DAT:AAQ45183
seq_documentation_block:
ID_AAQ45183 standard; cDNA to mRNA; 2533 BP.
XX
AC AAQ45183;
XX
XX 16-NOV-1994 (first entry)
XX
DE Soluble rice starch synthetic enzyme.
XX
KW Rice; starch synthetic enzyme; transit peptide; amyloplast;
XX transition; proloplast; expression; ss.
XX
OS Oryza sativa.
XX
XX Key Location/Qualifiers
XX 5'UTR 1..113
XX /*tag= a
FT
```

```
FT CDS 114..1994
FT /*tag= b
FT sig_peptide 114..452
FT /*tag= c
FT mat_peptide 453..1991
FT /*tag= d
FT 3'UTR 1992..2533
FT /*tag= e
XX
PN JP06070779-A.
XX
XX 15-MAR-1994..
XX
XX 07-JUL-1992; 92JP-0179947.
XX
XX 07-JUL-1992; 92JP-0179947.
XX
PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
XX
XX WPI; 1994-128678/16.
XX
XX P-PSDB; AAR51231.
XX
XX Soluble rice starch synthetic enzyme gene and transit peptide -
XX for the efficient transport of heterologous proteins to
XX amyloplast
XX
XX Claim 1; Page 15-17; 18pp; Japanese.
XX
XX The soluble rice starch synthetic enzyme gene has, at the N-terminal,
XX the transit peptide that is required for the transition of this
XX enzyme to the amyloplast. Introduction of this gene into the rice
XX proloplast augments the expression of soluble rice starch synthetic
XX enzyme. The transit peptide coding sequence can be used for
XX the efficient transition of any protein into amyloplasts.
XX
XX Sequence 2533 BP; 630 A; 550 C; 701 G; 652 T; 0 other;
SQ
alignment_scores:
Quality: 2821.50 Length: 726
Ratio: 4.522 Gaps: 18
Percent Similarity: 85.950 Percent Identity: 76.309
alignment_block:
US-09-674-824-2 x AAQ45183 ..
Align seg 1/1 to: AAQ45183 from: 1 to: 2533
2 AlaAlaThrGlyValGlyAlaGlyCysLeu...AlaProSerValArgLe 17
126 GCGGGGATGGGATCGGGCGCGCTGCTGTGGCGCGCGAGGTGAGG... 173
17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
174 .....CCGGGGAGG.....AGGT 186
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
187 TCGCGCTCCAGCGGTGGGAGG...CGGTGCGTGGCGGAGCTGAGCAGG 233
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaPro..... 65
234 GACGGTGGGTGGCG.....CACGGCGCGCTGGCACCGCGCGCC 271
66 .....ProLeuValProGlyPheLeuAlaPro.....ProP 76
272 GCTGTGTAAGCAGCGGCTCTCCGACCTTCTCGTGGCGAGCTGAGCGC 321
76 roProAlaProAlaGlnSer.....ProAlaProThrClnProLeu 90
322 CACCGCGCGCGCGAGTGGCGCGCGCGCGCGCGCGCGCGCGGTG 371
91 ProAspAlaGlyValGlyGluLeuAlaProAspLeuLeuGluGlyI 107
```



```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
372 CCGGACTCCGGCTGGGGAGATCGAGCCGAT.....CTAGAAGTCT 415
107 eAlaGluAspSerIleAspSerIleIleValAlaIleValAlaSerGluGlnAspS 124
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
416 CACAGAAGATTCCATCGCAAAACAATTTTGTGGCTAGTAGCAGGAGT 465
124 erGluIleMetAspAlaAsnGluGlnProGlnAlaAlaLysValThrArgSer 140
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
466 CTGAGATCATGGATGTGAAGGACCAAGCTCAAGCTAAAGTAACACCAGC 515
141 IleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLe 157
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
516 GTTGTCTTTGTACCCGGTGAAGCTTCTCCTTATGCAAGTCAGTGGACT 565
157 uGlyAspValCysGlySerLeuProIleAlaLeuAlaIleValGlyHis 174
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
566 AGGAGATGTTGGTGGTCTACTGCCAATGCTCTTGTCTTCTGGTGGTCATC 615
174 rgValMetValValMetProArgTyrLeuAsnGlySerSerAspLysAsn 190
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
616 GTCTGATGTTGTAATGCCGAGATACATGAACGGGSCCTTGAACAAAAAT 665
191 TyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheG 207
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
666 TTGCAAAACGCATTTTACACTGAGAAAGCACATTAAGATTCCATGCTTTGG 715
207 yGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValAspT 224
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
716 CGGAGAACATGAAGTACTTTTTTCCAGGATATAGGGATTCTGTGATT 765
224 rpValPheValAspHisProSerTyrHisArgProGlySerLeuTyrGly 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
766 GGGTGTGTTGTATCATCTCCTCATATCATAGACCTTGGAAATTTGATGGA 815
241 AspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCy 257
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
816 GATAATTTTGGTGCTTTTGGCGATAATAGTTTCAGATACACACTCCTGTG 865
257 sTyrAlaAlaCysGluAlaProIleIleLeuGluLeuGlyGlyTyrIleT 274
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
866 CTATCGGGCGCTGAAGCCCCATTAACTTGAACCTGGGAGATATATCT 915
274 yGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuVal 290
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
916 ATGGACAGAAATGCATGTTTGTGTGAATGATTGGCATGCCAGTCTGTG 965
291 ProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSe 307
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
966 CCAGTCTTCTTGCTGCAAAATATAGACCATATGGTGTATCAGGGATGC 1015
307 rArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProA 324
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1016 CCGCAGTGTCTTGTCATACATAATCTAGCACATCAGGCTGGGAGCCTG 1065
324 IaSerThrTyrProAspLeuGlyLeuProGluTyrTyrGlyAlaLeu 340
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1066 CCAGTACATATCTGACCTGGGATGCCACCTGAATGGTATGGAGCATTA 1115
341 GluTyrValPheProGluTrpAlaArgArgHisAlaLeuAspLysGlyGl 357
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1116 GAATGGGTGTTCCAGAGTGGCAGGGCGCATGCCCCCTTGAACAAGGGTGA 1165
357 uAlaValAsnPheLeuLysGlyValAlaValThrAlaAspArgIleValT 374
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1166 GCGAGTCAATTTTAAAGGCGCAGTGTGACAGCATCGAATGTGA 1215
374 hrValSerGlnGlyTyrSerTrpGluValThrThralaGluGlyGln 390
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1216 CTGTGAGCCAGGGGTATTTCATGGGAGGTCAACAACCTGCTGAAGSTGGCAA 1265
391 GlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleVa 407
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

1266 GGCTCAATCAGCTTTAAGCTCCCGAAGAGTGTATTGAATGAATTGT 1315
407 lAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuP 424
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1316 AAATGAATTTGACATTAATGATTGAACCCATCCACAGACAAGTTTCTCC 1365
424 roHisHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysAla 440
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1366 CTTATCATTTATCTGTGTGATGACCTGTCCGGAAGGCCAAGTGTAAAGT 1415
441 GluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuI 457
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1416 GAATTCAGAAGGAGCTGGSTTTACCTATAAGCCCGCATGTGCTCTGAT 1465
457 eGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysM 474
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1466 TGGCTTTATTGGAAAGATTGGACTATCAAAAGGCATTGATCTAATTAAC 1515
474 etAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuGly 490
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1516 TTGCCATTCCAGATCTCATCGGGACAATATTCAATTGTCATGCTTGA 1565
491 SerGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSerty 507
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1566 TCTGTGACCCAGGTTTGAAGGATGGATGAGATCCACAGAATCAGGGTA 1615
507 rLysAspLysPheArgGlyTrpValGlyPheSerValProValSerHis 524
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1616 CAGGGATAAATTTCTGGGATGGGTGGATTAGTTTCCAGATTTCACACC 1665
524 rgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluPro 540
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1666 GAATACTCCAGGTCGGATATATTGTTGATGCCATCCAGATTCGAACCT 1715
541 CysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValVa 557
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1716 TGTGGCTCAATCAGCTATATGCTATGCAATATGTTACAGTGCCTGTTGT 1765
557 lHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnPropheG 574
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1766 TATGGAATCGAGGCTCAGAGATACAGTGGAGAATTTTAACCCGTTTG 1815
574 lYalalysGlyGluGlyThrGlyTrpAlaPheSerProLeuThrVal 590
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1816 CTGAGAAAGGAGAGCAGGTCACAGGTGGSCATCTCTGCCACTAACCAIT 1865
591 Asp.LysMetLeuTrpAlaIleuArgThrAlaMetSerThrPheArgGluH 607
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1866 GAAAAAATGCTGTGGGCATTCGGGATGGCAATTTTCAGATACAGGGAAC 1915
607 lSlyProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHis 623
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1916 ACAAGTCTCTTGGGAGGCTTAATGCAAGCGAGGCATGCAAGCGACTTT 1965
624 ThrTrpAspHisAla.ProSerSerThrSerArgSerSerSerGlyProS 640
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1966 ACATGGACCATGCCGCTCAGAGTATGAACAGATCTTCGAATGGSCCTT 2015
640 erTrpThrAsnProThrSerCysArg.ArgGlyLeuGlyArgSerLys.. 655
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2016 CATGATCAACCATATGCTGTAATGATTTGAAGGAAGCAGCGCAATT 2065
656 .CysGluSerPro.....SerAlaLeuLysThrS 665
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2066 TCTCCGAGGACCTCAATCTTCTGCTTTTCATGCGGAATGAAACATT 2115
665 erSerSerSerPhe.ArgGly.....Pr 672
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2116 TGTACACTACATGGAAGGAGCAACCACTTATGCAAAAGTTGCAACGATCAC 2165
672 oGluGlyTyrProCysThrLeuArgCysProAlaThrValGluSerGlnC 689
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2166 TCAAGGTTACCTTGTAGG.....CCTGCTACTTGGCCA..... 2199

```



689 ysalacysLeuLeuTrpPhe 695  
 2200 .....ATATGGTTC 2208

seq\_name: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT:AAT95785

seq\_documentation\_block:  
 ID AAT95785 standard; cDNA to mRNA; 2383 BP.

AC AAT95785;

DT 22-MAY-1998 (first entry)

DE Maize starch synthase type I cDNA.

KW Maize; starch synthase type I; starch; ds.

OS Zea mays.

Key Location/Qualifiers

FT CDS 2..1951

FT /tag- a

FT /product= starch\_synthase\_type\_I

PN DE19619918-A1.

XX 20-NOV-1997.

XX 17-MAY-1996; 96DE-1019918.

XX 17-MAY-1996; 96DE-1019918.

XX (PLAN-) PLANTTEC BIOTECNOLOGIE GMBH.

XX Froberg C, Kossmann J;

XX WPI; 1998-000821/01.

DR P-PSDB; AAW38218.

XX DNA encoding maize starch synthase type I protein - for producing

XX transgenic plants

XX Claim 1: Pages 16-20; 23pp; German.

XX The present sequence encodes maize starch synthase type I, useful  
 CC in the production of starch. Starch can be used in various  
 CC conventional starch applications, e.g. starch hydrolysate products,  
 CC foods, papermaking, adhesives, textiles, building materials, soil  
 CC stabilisation, agrochemicals, pharmaceuticals, cosmetics, coal  
 CC briquettes, ore and coal slurries, foundry casting, rubber, leather  
 CC and synthetic polymers. The enzyme produces a starch stated to have  
 CC different physicochemical properties, especially viscosity and  
 CC gelling properties, from wild type starch.

SQ Sequence 2383 BP; 582 A; 546 C; 633 G; 622 T; 0 other;

alignment\_scores:

Quality: 2809.50 Length: 749  
 Ratio: 4.411 Gaps: 10  
 Percent Similarity: 85.047 Percent Identity: 73.565

alignment\_block:

US-09-674-824-2 x AAT95785 ..

Align seg 1/1 to: AAT95785 from: 1 to: 2383

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20

41 TCGCCGCGTGGCGCGCGTCCCTC..... 67

20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34

```

 68 .CTCGCGCGCGCGCGCTGGCGCGCGCGCGCTGGCGGACCGCGCGCGCGCGC 116
 35 ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
 117 GCGCGCTCCAGCGCGTCTGCGCGCGCGTGGTGGCGGAGCTGAGCAGG 166
 51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
 167 GAGGCGCGCGCGCGCGTCCACCGCGCGTCTGGCGCGCGCGCT 216
 67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84
 217 CGTGGCGCGCTCTCGCG...CGCGCGCGCGAGCCGCCGCGGTGAGCGG 263
 84 laProThrGlnProProProProAlaGlyValGlyGluLeuAlaPro 100
 264 CATTGAGCGCGCGCGCGTGGCGCGCGCGTGGGGTCTCGGTGTC 313
 101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117
 314 GAA.....CCTGAAGGATTGCTGAAGGTTCCATCGATAACACAGTAGT 357
 117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
 358 TGTGGCAAGTGAGCAAGATTCTGAGATTGTGGTGGAAAGGAGCAAGCTC 407
 134 lnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
 408 GAGCTAAAGTAAACACAAACATTGCTCTTTGTAAGTGGCGAAGCTTCTCT 457
 151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
 458 TATGCAAAAGCTGGGGGTCTAGGAGATGTTTGTGGTTCATTGGCAGTTGC 507
 167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
 508 TCTTGTCTGCTGGTGCACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 557
 184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
 558 ATGTGTACCTCCGATAAGAAATTATGCAAAATGCAATTTACACAGAAAAAC 607
 201 lLelysIleProCysPheGlyGlySerHisGluValThrPhePheHisGl 217
 608 ATTCGGATTCCATGCTTTGGCGGTGAACATGAAGTTACCTTCTTCCATGA 657
 217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
 658 GTATAGAGATTCAGTTCAGTGGGTGGTGGTGGTGGTGGTGGTGGTGGT 707
 234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
 708 GACCTGGAATTTATATGGAGATAAGTTGGTGGTGGTGGTGGTGGTGGT 757
 251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
 758 TTCAGATACACACTCCTTGTCTATGCTGCATGTGAGGCTCTTTGGTCT 807
 267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
 808 TGAATTGGGAGGATATATTATGGACAGAAATTGCATGTTTGTGTCGAAT 857
 284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
 858 ATTGGCATGCCAGTCTAGTGCAGTCTCTTGTGTCGCAAAATATAGACCA 907
 301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
 908 TATGGTGTATTATAAGACTCCCGCAGCATCTTGTGTAACATATATTAGC 957
 317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334
  
```



```

958 ACATCAGGGTGTAGAGCTGCAAGCACATATCCTGACCTTGGGTGCCAC 1007
334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaIleArgArg 350
|||||
1008 CTGAATGGTATGGAGCTCTGGAGTGGGTATTCCTCGAATGGGGAGGAGG 1057
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
|||||
1058 CATGCCCTTGACAAGGGTGAGCGAGTTAAUTTTTTGAAGGTCAGCTGT 1107
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
|||||
1108 GACAGCAGATCGAATCGTGAATGTCAGTAAGGGTTATTCATGGGAGGTCA 1157
384 hrThrAlaGluGlyGlnGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
|||||
1158 CAACTGCTGAAGGTGGAGCGGCTCAATGAGCTCTTAAGCTCCAGAAAG 1207
401 SerValLeuAsnGlnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
|||||
1208 AGTGATATTAAACGGAATTGTAATGGAATTGACATTAATGATTGAACCC 1257
417 oThrThrAspLysCysLeuProHisLysTyrSerValAspAspLeuSerG 434
|||||
1258 TGCCACAGACAATGATCCCTGCTCATTTATCTGTGATGACCTCTCTG 1307
434 lyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
|||||
1308 GAAAGGCCAAATGTAAGGTGCATCTCAGAAAGGAGCTGGGTATACCTATA 1357
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
|||||
1358 AGCCCTGATGTCCTCTGATTTGGCTTTATTTGGAAGATTGGATTATCAGAA 1407
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
|||||
1408 AGGCATTGATCTCATTCACATTTATCATACCATCATCTATCGCGGAAGATG 1457
484 aGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
|||||
1458 TTCAATTTGTCATGCTTGGATCTGCTGTGACCCAGAGCTGAAGATTGGATG 1507
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
|||||
1508 AGATCTACAGATCGATCTTCAAGGATAAATTCGTGGATGGGTGGATT 1557
517 eSerValProValSerHisArgIlePheThrAlaGlyCysAspIleLeuLeuM 534
|||||
1558 TAGTGTTCAGTTTCCACCGAATRAACTGCCGCTCGCATATATTGTTAA 1607
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
|||||
1608 TGCCATCCAGATTCGAACCTTGTGCTCTCAATCAGCTATATGCTATGCAG 1657
551 TyrGlyThrValProValIleHisGlyThrGlyGlyLeuArgAspThrVa 567
|||||
1658 TATGACAGATTCCTGTGTCATCCACTGGGGCCCTTAGAGATACCCGT 1707
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpA 584
|||||
1708 GGAGAACTTCAACCCCTTCGTGTGAGATGGAGAGCAGGTTACAGGTGGG 1757
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
|||||
1758 CATTCGACCCCTTAACCAAGAAACATGTTGTGGACATTCGGAACCTGCA 1807
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
|||||
1808 ATATCTACATACAGGGAACACAAAGTCTCTCTGGGAAGGCTTAATGAACGG 1857
617 gGlyMetThrLysAspHisThrTrpAspHisAlaProSerSerThrSer 633
|||||
1858 AGGCATGCAAAAGACTTCACGTGGGACCATCGCGCTGAACAAATACGAAC 1907

```

```

634 ArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG1 650
|||||
1908 AAATCTTCCAGTGGGCTTTCATCGATCGACCTTATGTCATGTAATAAAGG 1957
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSer 666
|||||
1958 A.....CCAAAGTGGTGTCTTGAAGATCATCAGTTTCATCATCC 1998
667 SerSerPheArgGlyProGluGlyTyrPro.....CysThrLeuArgCy 681
|||||
1999 TATAGTAAGCTGAATGATGAAGAAACCCCTGTACATTACATGGAAGGC 2048
681 sProAlaThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlys 698
|||||
2049 AGACCGGCTATTGGCTCCATTGCTCCAATGCTGCTTTGGCTGCCT... 2094
698 erArgThrTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGly 714
|||||
2095 .....TGCCTCGATGGACCGGATCGAGTGAGGAATCCA 2127
715 ArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAla.....Al 728
|||||
2128 GCCGAACGACAGTGTGAAGGATAGGAAGGGAGCTGGAAGCAGTCACGC 2177
728 aGlyTrpLeuThrAlaLysHisSerAspGlySerLeuSer 742
|||||
2178 AGGCAGCTCGCCGTGATTCATATGGAACAAGCTGGAGTCAGT 2220

```

seq\_name: /SIBS1/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50643

seq\_documentation\_block:

ID\_AAZ50643 standard; cDNA; 2008 BP.

AC\_AAZ50643;

23-MAY-2000 (first entry)

Corn soluble starch synthase gene fragment inserted in pSS65-C11.

Soluble starch synthase; starch fine structure; corn;  
transgenic plant; amylose; amylopectin; amylose polymerisation;  
non-granule bound starch synthase; non-GBSSI; altered starch; food;  
paper; plastic; adhesive; ss.

Zea mays.

WO200006755-A2.

10-FEB-2000.

26-JUL-1999; 99WO-US16296.

28-JUL-1998; 98US-0094436.

(DUPO ) DU PONT DE NEMOURS & CO E I.

Broglie KE, Lightner JE;

WPI; 2000-195311/17.

Producing transgenic cereal crops with altered starch structure useful  
for preparing foodstuff, paper, plastic or adhesives, comprises  
transforming crops with chimeric sense or antisense gene construct  
encoding starch synthase

Claim 5; Page 53; 56pp; English.

The present sequence is the corn soluble starch synthase (SSI) DNA

sequence comprising the entire SSI coding region and a 3' UTR fragment

inserted into plasmid pSS65-C11. The chimeric gene containing the zein

promoter followed by the 3'UTR is used as a sense construct for

preparation of transgenic corn expressing altered starch structure. The



217 uTyrArqAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234  
|||||:::|||||  
631 GTATAGAGATTCACTGGGTGTTGTCATCATCCCTCATATCAC 680

234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250  
|||||:::|||||  
681 GACCTGSAAATTTATATGGAGATAAGTTGGTGCTTTGGTCATAATCAG 730

251 pheArgTyrThrLeuLeuCysTyrAlaAlaLacCysGluAlaProLeuIleLe 267

251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe

267 uCluLeuGlyGlyTyrIleTyrGlyClnAsnCysMetPheValValAsnA 284  
 781 TGAATTGGGAGGATATATTATGGACAGAAATGCATGTTGTGTCAATG 830

284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300

831 ATGGCATGCCAGTCTAGTGCCAGTCCTTCTTGCTGCAAAATATAGACCA 880

[illegible]

881 TATGGTGTTTATAAGACTCCCGCAGCATTCCTGTGAATACATAATTAGC 930

317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuPro 334

931 ACATCAGGGTGTAGAGCCTGCAAGCACATATCCTGACCTTGGGTTGCCAC 980

334 roGluTrpTyrG[V]aIaLeuGluTrpValPheProGluTrpAlaArqArq 350

[illegible]

981 CIGATGGTATGGAAGACCTCGAAGTCGGGTTTCCCTCGATTGCCGCGCAGCAAT

351 HisAlaLeuAspLysGlyGluAlaValAsnPhenLeuLysGlyAlaVal 367

1031 CATGCCCTTGACAAGGGTGAGGCAGTTAATTTTGAAGCTGCAGTTGT 1080

367 IThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384

1081 GACAGCAGATCGAATCGTGACTGTCAGTAAGGGTTATTCTGGGAGGTCA 1130

384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400

1131 CAACTGCTGAAGGTGGACAGGGCCTCAATGAGCTCTTAAGCTCCAGAAAG 1180

401 `serval[leu]asnG[v]I[eva]AsnG[v]I[eAsp]I[eAsp]trpAsnPr 417`

401  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
191

1181 AGTGTATTAAACGGAA11G1ATAA1GGAA11GACAA1TAATG1ATGCGG11CC 1239

```

417 oThrAspLysCysLeuProHisHistyrSerValAspPleureg 434
|::|||:|||::||| ||| ||| ||| ||| ||| ||| |||
|::|||:|||::||| ||| ||| ||| ||| ||| ||| |||

```

1231 TGGCACAGACAAATGTATCCCCCTGTATTCTGTTGATGACCTCTCTG 1280

434 lyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProval 450

1281 GAAAGGCCAATGTAAAGGTGCATTGCAGAACGAGCTGGGTTTACCTATA 1330

451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467

1331 AGCCCTGATGTTCCCTCTGATTGGCTTTATTGGAAGGTTGGATTATCAGAA 1380

1557 sclytLeasntentIeiysMetAlaIeprogIuleuMetArgGluAspV 484

[illegible]

1381 AGGCATTGATCTCATTCAACTTATCATACCAAGATCAGATCCGCGGAGGAGGATG 1430

484 aGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500

1431 TTCAATTGTGTCATGCTTGGATCTGGTGACCCAGAGCTTGAAGATTGGATG 1480

501 ArgSer<sup>Thr</sup>GluSerSer<sup>Tyr</sup>LysAspLysPheArgGlyTrpValGlyPh 517

1481 AGATCTACAGATCGATCTTCAAGGATAAAATTCGTGGATGGGTTGGATT 1530



```

517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
|||||
1531 TAGTCTTCAGTTCCACCACCAATAAATCGCGGCTGCATATATGTTAA 1580
.
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
|||||
1581 TGCCATCCAGATTGGAACCTTGTGGTCTCAATCAAGTATATGCTATGCAG 1630
.
551 TyrGlyThrValProValHisGlyThrGlyLeuArgAspThrVa 567
|||||
1631 TATGACACAGTCTCTGTGTCCATGCACACTGGGGCCCTTACAGATACCGT 1680
.
567 lGluThrPheAsnPropheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584
|||||
1681 GGAGAACTTCAACCCCTTTCGGTGAGAAATGGAGACGAGCGGTACAGGTGG 1730
.
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
|||||
1731 CATTCGCACCCCTAACACACAGAAACATGTTGTGGACATTGCGAACTGCA 1780
.
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
|||||
1781 ATATCTACATACAGGAACACAACTCTCTCTGGGAAGGCTATATGAAGCG 1830
.
617 gGlyMetThrLysAspHisThrTrpAspHisAlaPro. SerSerThrSer 633
|||||
1831 AGGCATGTCAAAAGACTTCACGTGGGACCATCGCGCTGAACAATACGAAC 1880
.
634 ArgSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG1 650
|||||
1881 AAATCTTCCAGTGGGCTTCATCATCGACCCCTATGTCATGTATAAAAAA... 1928
.
650 yLeuGlyArgSerLysCysGluSerPro 659
|||||
1929 .....AGGACCAAGTGTGTCTCT 1949

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV29756
seq_documentation_block:
ID AAV29756 standard; DNA; 1752 BP.
XX
AC AAV29756;
XX
DT 11-SEP-1998 (first entry)
XX
DE Zea mays soluble starch synthase I gene.
XX
KW SER; starch-encapsulating region; fusion vector;
KW soluble starch synthase I; glucosyl transferase; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 1..2007
FT FT /*tag= a
FT FT /product= soluble starch synthase I
XX
PN W09814601-A1.
XX
PD 09-APR-1998.
XX
PF 30-SEP-1997; 97MO-US17555.
XX
PR 30-SEP-1996; 96US-0026855.
XX
PA (EXSE-) EXSEED GENETICS LLC.
XX
PI Guan H, Keeling P;
XX
DR WPI: 1998-240100/21.
XX
DR P-PSDB; AAW56488.
XX
PT Hybrid polypeptide comprising starch-encapsulating region and

```

```

PT protein - useful for, e.g. producing protein(s) resistant to
XX degradation by stomach acids
XX
XX Example 2; Page 39-41; 156pp; English.
XX
XX The sequence is that of the soluble starch synthase I gene.
XX It can be used in the production of a hybrid polypeptide
XX comprising a starch-encapsulating region (SER) fused
XX to a payload protein. The hybrid polypeptide can be used to make
XX modified starches comprising the payload protein, selected from,
XX e.g. hormones, growth factors, antibodies, enzymes, dyes,
XX immunoglobulins, etc. The modified starch can also be used
XX to provide grain feeds enriched in amino acids. By encapsulating
XX the payload protein in starch, it is more resistant to
XX degradation by stomach acids.
XX
XX Sequence 1752 BP; 433 A; 379 C; 469 G; 471 T; 0 other;

```

```

alignment_scores:
Quality: 2709.50 Length: 585
Ratio: 4.909 Gaps: 3
Percent Similarity: 94.359 Percent Identity: 86.154

alignment_block:
US-09-674-824-2 x AAV29756 ..

Align seg 1/1 to: AAV29756 from: 1 to: 1752

45 ValAlaGluLeuSerArgGluGlyProAlaAlaArgProAlaGlnGlnG1 61
4 GTCCGGAGCTGACAGAGGGGGCCGCCGCCGCCGCTGCCACCCGC 53
61 nGlnLeuAlaProProLeuValProGlyPheLeuAlaProProProA 78
54 GCTGCTGGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 100
78 laProAlaGlnSerProAlaProThrGlnProProLeuProAlaGly 94
101 AGCCACACGGGTGAGCGGCATCGACGCCGCCGCCGCCGCCGCCGC 150
95 ValGlyGluLeuAlaProAspLeuLeuGluGlyLeuAlaGluAspSe 111
151 CTGGGGACCTCGGT.....CTCGAACCTGAAGGATTGCTGAAGTTC 194
111 rIleAspSerIleIleValAlaAlaSerGluGlnAspSerGluIleMe 128
195 CATCGATAACACAGTAGTTGTGGCAAGTGAAGCAAGATTCTGAGATT 244
128 spAlaAsnGluGlnProGlnAlaLysValThrArgSerIleValPhe 144
245 TTGGAAGAGGAGGACGCTCGAGCTAAGTAACACAAAGCATTGTCTT 294
145 ThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuGlyAspVal 161
295 ACCGGCAAGCTTCTCTTATGCAAGTCTGGGGTCTAGGAGATTGTTG 344
161 sGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMet 178
345 TGGTTCAATTGCCAGTTGCTTTCGCTGCTGCTGCTGCTGCTGCTG 394
178 alMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLys 194
395 TAATGCCAGATATTTAATGCTACCTCCGATAAGAATTATGCAATGCA 444
195 LeuTyrThrAlaLysHisIleLysIleProCysPheGlyGlySerHis 211
445 TTTTACAGAAAAACACATTCGATTCATGCTTTGGCGGTGAACATGA 494
211 uValThrPhePheHisGluTyrArgAspAsnValAspTrpValPhe 228
495 AGTTACCTTCTTCCATGATAGAGATTACGTTGCTGCTGCTGCTGCT 544

```



228 sPHsProSerTyrHisArgProGlySerLeuTyrGlyAspAsnGly 244  
 545 ATCATCCTCATATCAAGACCTGGAAATTTATATGAGATAAGTTGGT 594  
 245 AlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCystTyrAlaAla 261  
 595 GCTTTTGGTGATAATCATGTTTCAGATACACACTCCTTTGCTATGCTGCATG 644  
 261 sGluAlaProLeuLeuLeuLeuGlyGlyTyrIleTyrGlyGlnAsn 278  
 645 TGAGGCTCCTTTGATCCTTGAATGGGAGGATATATTTATGGACAGAATT 694  
 278 yMetPheValValAsnAspTrpHisAlaSerLeuValProValLeuLeu 294  
 695 GCATGTTTGTTCATGATGATGGCATGCCAGTCTAGTGCAGTCTTCCT 744  
 295 AlaAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLe 311  
 745 GCTGCAAAATATAGCATATGTTGTTTATAAAGACTCCCGCAGCATCT 794  
 311 uValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrP 328  
 795 TGTAAATACATAATTTAGCACATCAGGGTGTAGACCTGCAGACACATATC 844  
 328 roAspLeuGlyLeuProProGluTyrTyrGlyAlaLeuGluTyrValPhe 344  
 845 CTGACCTTGGCTTCCACCTGAATGGTATGGAGCTCTGGAGTGGGTATTC 894  
 345 ProGluTyrPAlaArgArgHisAlaLeuAspLysGlyGluAlaValAsnPh 361  
 895 CTTGAAGTGGTGGAGGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 944  
 361 eLeuLysGlyAlaValValThrAlaAspArgIleValThrValSerGlnG 378  
 945 TTTGAAGTGGTGGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1044  
 378 lyTyrSerTrpGluValThrThrAlaGluGlyGlyGlnGlyLeuAsnGlu 394  
 995 GTTATTCCTGGAGGTCACACTGCTGAAGCTGGACAGGCTCAATGAG 1044  
 395 LeuLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAs 411  
 1045 CTCTTAAGCTCCAGAAAGAGTGTATTAACCGGAATGTAAATGGAATGA 1094  
 411 pIleAsnAspTrpAsnProThrThrAspLysCysLeuProHisHisTyrS 428  
 1095 CATTAATGATGGAACCTTCCACAGACAATGATCCCTGTCATTATT 1144  
 428 erValAspAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGlnLys 444  
 1145 CTGTTGATGACCTCTCTGAAAGGCCAAATGTAAAGGTGCATTGCAAG 1194  
 445 GluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheIleG 461  
 1195 GAGCTGGGTATTACCTATAGGCTCATGTTCTCTGATTTGGCTTTATGG 1244  
 461 yArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIleProG 478  
 1245 AAGTTGGATATACAGAAAGGCAATGATCATCAACTTATCATACACAG 1294  
 478 luLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAspPro 494  
 1295 ATCTCATGCGGGAAGATGTTCAATTTGTCATGCTTGGATCTGGTGACCCA 1344  
 495 IlePheGluGlyTyrPheMetArgSerThrGluSerSerTyrLysAspLysPh 511  
 1345 GAGCTTGAAGATGGATGAGATCTACAGATGATCTTCAAGGATTAAT 1394  
 511 eArgGlyTyrPValGlyPheSerValProValSerHisArgIleThrAlaG 528  
 1395 TCGTGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1444  
 528 lyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLeuAsn 544

1445 GCTGCGATATATTGTTAATGCCATCCAGATTGCAACCTTGTGCTCTCAAT 1494  
 545 GlnLeuTyrAlaMetGlnTyrGlyThrValProValValHisGlyThrG 561  
 1495 CAGCTATATGCTATGAGATGCGATGCGACAGTTCCTGTTGTCATGCAACTGG 1544  
 561 yGlyLeuArgAspThrValGluThrPheAsnProPheGlyAlaLysGlyG 578  
 1545 GGCCTTTAGAGATACCGTGGAGAACTTCAACCTTTCCGTTGAGAAATGGAG 1594  
 578 luGluGlyThrGlyTyrAlaPheSerProLeuThrValAspLysMet.Le 594  
 1595 AGCAGGGTACAGGGTGGCATTCGCCACCTTAACCAACACAAAAACATGTTT 1644  
 594 uTTPAlaLeuArgThrAlaMetSerThrPheArgGluHisLysProSerT 611  
 1645 CTGACATTCGCAACTGCAATATCTACATACAGGGAACACAAAGTCTCCT 1694  
 611 rpGluGlyLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHis 627  
 1695 GCGAAGGCTAATGAAGCGAGGATGTCAAAGACTTCACGTGGGACCAT 1744  
 628 Ala 628  
 1745 GCC 1747  
 seq\_name: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT:AAV70960  
 seq\_documentation\_block:  
 ID AAV70960 standard; DNA; 1749 BP.  
 XX  
 AC AAV70960:  
 XX  
 DT 23-AUG-1999 (first entry)  
 XX  
 DE DNA encoding maize starch soluble synthase 1-2..  
 XX  
 KW Non-glycogen-like polysaccharide production; fermentation;  
 KW starch synthesis enzyme; glycogen-synthesis enzyme; glycogen synthesis;  
 KW non-starch branching gene; amylopectin; amylose; plant-like starch;  
 KW maize starch soluble synthase 1-2; ss.  
 XX  
 OS Zea mays.  
 XX  
 PN W09844780-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 03-APR-1998; 98WO-US06660.  
 XX  
 PR 04-APR-1997; 97US-0042939.  
 XX  
 PA (EXSE-) EXSEED GENETICS LLC.  
 XX  
 PI Guan H, Keeling PL;  
 XX  
 DR WPI; 1998-568285/48.  
 DR P-PSDB; AAV70894.  
 XX  
 PT Producing non-glycogen-like polysaccharides in bacteria, fungi or  
 PT plants - transformed with genes for enzymes involved in starch or  
 PT glycogen synthesis allows fermentative production of starches with  
 PT engineered properties  
 PS  
 PS Disclosure; Fig 52; 150pp; English.  
 CC  
 CC The specification describes a method for the production of  
 CC non-glycogen-like polysaccharides in a host. The method comprises  
 CC transforming a host, suitable for fermentation, with genes encoding  
 CC starch- or glycogen-synthesis enzymes, and fermenting the transformants.  
 CC The specification also describes hosts transformed with a gene active  
 CC in glycogen synthesis and at least one non-starch branching gene.



CC involved in production of amylopectin or amylose in its original host.  
 CC The method is used to produce plant-like starches by fermentation and  
 CC new starches in plants. These starches are useful for all food and  
 CC non-food applications of starch. The present sequence is used in  
 CC the course of the invention.

XX  
 S0 Sequence 1749 BP: 432 A; 378 C; 469 G; 470 T; 0 other;

## alignment\_scores:

Quality: 2708.50 Length: 585  
 Ratio: 4.916 Gaps: 3  
 Percent Similarity: 94.188 Percent Identity: 86.154

## alignment\_block:

US-09-674-824-2 x AAV70960 ..

Align seg 1/1 to: AAV70960 from: 1 to: 1749

```

45 ValAlaGluSerArgGluGlyProAlaAlaArgProAlaGlnGlnI 61
|||||
4 GTCGCGGAGCTGAGCAGGAGGGCGCGCGCGCTGCCACCGC 53
|||||
61 nGlnLeuAlaProLeuValProGlyPheLeuAlaProProA 78
|||||
54 GCTGCTGGCGCGCGCTGCTGCGCGCTTCTCGCG...CCGCGCGCG 100
|||||
78 laProAlaGlnSerProAlaProThrGlnProLeuProAspAlaGly 94
|||||
101 AGCCACCGGGTGAGCGCGCTGACGCGCGCGCGCGCGCGCGCGCG 150
|||||
95 ValGlyGluLeuAlaProAspLeuLeuGluGlyIleAlaGluAsp 111
|||||
151 CTGGGGACCTCGGT.....CTCGAACCTGAAGGATTCGTAAGCTTC 194
|||||
111 rIleAspSerIleIleValAlaAlaSerGluGlnAspSerGluIle 128
|||||
195 CATCGATAACACAGTAGTTGGCAAGTGAGCAAGATTCGAGATTGTG 244
|||||
128 spAlaAsnGluGlnProGlnAlaLysValThrArgSerIleValPhe 144
|||||
245 TTGGAAGGAGCAAGCTCGAGCTAAAGTAAACACAAACCAATGCTTT 294
|||||
145 ThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuGlyAspVal 161
|||||
295 ACCGCGGAGCTCTCTCTATGCAAGCTCGGGGCTCTAGAGATTGTTG 344
|||||
161 sGlySerLeuProIleAlaLeuAlaArgGlyHisArgValMetVal 178
|||||
345 TGGTTATTGCCAGTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 394
|||||
178 alMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLys 194
|||||
395 TAATGCCACAGATATTTAAATGGTACCTCCGATGAAGATTAAGCA 444
|||||
195 LeuTyrThrAlaLysHisIleLysIleProCysPheGlySerHisG 211
|||||
445 TTTTACACAAACACATTCGGATCCAGTCTTGGCGGTGAACATGA 494
|||||
211 uValThrPhePheHisGluTyrArgAspAsnValAspTrpValPhe 228
|||||
495 AGTTACCTCTTCATGAGTATAGAGATTGAGTTGACTGGGTGTTGTT 544
|||||
228 spHisProSerTyrHisArgProGlySerLeuTyrGlyAspAsnPhe 244
|||||
545 ATCATCCCTCATATACAGACCTGGAAATTTATATGAGATAAGTTGG 594
|||||
245 AlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAla 261
|||||
595 GCYTTTTGGTGATAATCAGTTTCAGATACACACACCTCTTGTGCTG 644
|||||
261 sGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGlnAs 278
|||||

```

```

645 TGACGCTCTTTGATCCTTGAATTGGGAGGATATATTTATGGACAGA 694
|||||
278 yMetPheValValAsnAspTrpHisAlaSerLeuValProValLeu 294
|||||
695 GCATGTTTGTTCATGATGGCATGCCAGCTAGTCCAGTCTCTCTCT 744
|||||
295 AlaAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThr 311
|||||
745 GCTGCAAAATATAGACCATATGGTGTATATAAAGACTCCCGCAGCAT 794
|||||
311 uValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyr 328
|||||
795 TGTAAATACATAATTTAGCACATCAGGCTGAGACCTGCAAGCACAT 844
|||||
328 roAspLeuGlyLeuProProGluTyrPyrGlyAlaLeuGluTyrVal 344
|||||
845 CTGACCTTGGTTGCCACCTGAATGGTATGGAGCTCTGGAGTGGGTAT 894
|||||
345 ProGluTrpAlaArgArgHisAlaLeuAspLysGlyGluAlaValAsn 361
|||||
895 CCTGAATGGCGGAGGAGGATGCCCTTGACAAGGGTGAGGAGTTAAT 944
|||||
361 eLeuLysGlyAlaValValThrAlaAspArgIleValThrValSer 378
|||||
945 TTTGAAAGGTGCAGTTGTGACACAGATCGATCGTGTGAGTGTGAG 994
|||||
378 lyTyrSerTrpGluValThrAlaGluGlyGlnGlyLeuAsnGlu 394
|||||
995 GTTATTCTGTGGAGGTACAACTGCTGAAGTGGACAGGGCCTCAATG 1044
|||||
395 LeuLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIle 411
|||||
1045 CTCTTAAGCTCCAGAGAGTGTATTAAACGGAATTTGTAATGGAAT 1094
|||||
411 pIleAsnAspTrpAsnProThrThrAspLysCysLeuProHisHisTyr 428
|||||
1095 CATTAATGATTGCAACCTGCCACACAGAAATGTATCCCTGTCATAT 1144
|||||
428 erValAspAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGln 444
|||||
1145 CTGTTGATGACCTCTCTGGAAGGCGCAATGTAAAGGTGCATTTGCA 1194
|||||
445 GluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheIle 461
|||||
1195 GAGCTGGGTTTACCTATAGGCTGATGTTCTCTGATTTGGCTTTAT 1244
|||||
461 yArgLeuAspTyrGlnLysGlyIleAspLeuLeuLysMetAlaIle 478
|||||
1245 AAGGTTGGATTATCAGAAAGGCAATTCATCTCAACTTATCATACCA 1294
|||||
478 luLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAsp 494
|||||
1295 ATCTCATGGGGAAGATGTTCAATTTGTCATGCTGGATCTGGTGACCA 1344
|||||
495 IlePheGluGlyTrpMetArgSerThrGluSerSerTyrLysAspLys 511
|||||
1345 GAGCTTGAAGATTTGATGAGATCTACAGAGTCGATCTTCAAGGATA 1394
|||||
511 eArgGlyTrpValGlyPheSerValProValSerHisArgIleThrAla 528
|||||
1395 TCGTGGATGGGTGGATTAGTTGTTCCAGTTTCCACCGCAATAACTGG 1444
|||||
528 lyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLeuAs 544
|||||
1445 GCTGCGATATATTTGTAATGCCATCCAGATTCGAACCTTGTGCTCA 1494
|||||
545 GlnLeuTyrAlaMetGlnTyrGlyThrValProValIleHisGlyThr 561
|||||
1495 CAGCTATATGCTATGAGTATGGCAGTCTCTGTTGTCTCACTCACTGG 1544
|||||
561 yGlyLeuArgAspThrValGluThrPheAsnProPheGlyAlaLysGly 578
|||||
1545 GGGCCTTAGAGATACCGTGGAGAACTTCAACCTTTTCGGTGAGATGG 1594
|||||

```



```

alignment_block:
  US-09-674-824-2 x AAT67285
  Align seg 1/1 to: AAT67285 from: 1 to: 2992
  4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  614 TCGGCGGTGGGCGCGCGTGCCTC.....CTCCTCGCGCG 648
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAlaArgLeuA 37
    :|||:||||:||||:||||:||||:||||:||||:||||:
  649 GNGCGCTTGGCGCGCGTGCAGCGACCGCGCGCGCGCGCGCGAGTCC 698
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  37 rgArgCLeuAlaArgGlyArgTyrValAlaGluLeuSerArgGluGlyPro 53
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  699 AGCGCGTGTGGCGCGCGGTGCTGCTCGCGAGACTGAGCAGGAGGGGCC 748
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  54 AlaAlaArgProAlaGlnGlnGlnLeuAlaProProLeuValProG1 70
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  749 GCGCGCGCGCGCTGCCACCGCGCGTGTGCGCGCGCGCTGCTGCGCGG 798
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  70 yPheLeuAlaProProProAlaProAlaGlnSerProAlaProThrG 87
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  799 CTTCTCTCGG...CGCGCGCGCGAGCCCGCGGTGCGCGCGAGTCGACG 845
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  87 InProProLeuProAspAlaGlyValGlyGluLeuAlaProAspLeuLeu 103
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  846 CGCGCGCGTGGCGCGCGCGCTGGGACCTCGGT.....CTCGAA 889
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  104 LeuGluGlyIleAlaGluAspSerIleAspSerIleIleValAlaLeu 120
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  890 CTTGAAGGGATTGCTGAAGTTTCCATTCGATACACAGTAGTTGTGGCAAG 939
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  120 rGluGlnAspSerGluIleMetAspAlaAsnGluGlnProGlnAlaLysV 137
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  940 TGAGCAGAGATTCTGAGATTGTGGTTGGAAGAGACAGCTCGAGCTAAG 989
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  137 alThrArgSerIleValPheValThrGlyGluAlaAlaPro..TyrAla 153
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  990 TAACAAAGACATTGTCTTTGTGAACGGCGAAGCTTCCTTTAATCGAA 1039
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  153 sSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuAlaA 170
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1040 GTCGTGGGGTCTAGGAGATGTTTGTGGTTCAATGCCAGTTGCTCTGCTG 1089
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  170 laArgGlyHisArgValMetValMetProArgTyrLeuAsnGlySer 186
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1090 CTCGCGCTCACCGTGTGATGGTTGTAATGCCAGACATTTAATGGTACC 1139
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  187 SerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLys1 203
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1140 TCCGATTAAGAATTATGCAATGCATTTTACTCAGAAAAACACATTCGGAT 1189
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  203 eProCysPheGlyGlySerHisGluValThrPhePheHisGluTyrArgA 220
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1190 TCCATTCTTTGGCGGTGAACATGAAGTTTACCTTCTTCCATGATGATAGAG 1239
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  220 sPasnValAspTrpValPheValAspHisProSerTyrHisArgProGly 236
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1240 ATTCACTGACTGGGTGTTTGTGATCATCCCTTCATATCAGACACCTGGA 1289
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  237 SerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTy 253
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1290 AATTATATGGACATAAGTTTGTGCTCTTTTGGTGATAATCAGTTCAGATA 1339
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  253 rThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuLeuLeu 270
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1340 CACACTCCTTGTGATGCTGCTATGTGAGGCTCCTTTGGTCCCTTGAATGG 1389
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  270 lyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrpHis 286
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1380 GAGCATATATTTATGACACAAATTCATGCTGTGCTTCATGATGATCGCAT 1439
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

578 luGlucGlyThrClyTrpAlaPheSerProLeuThrValaspLysMet..le 594
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1595 AGCAGGGTACAGGGTGGGCATTCGCACCCCTAACACACAGAAAACATGTTT 1644

594 utrPalaleuArgThrAlaMetSerThrPheAtqGluHisLysLysProSerT 611
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1645 GTGGACATTCGCAATATCTACATACAGGGAACACAGTCTCTCT 1694

611 rpGlucGlyLeuMetLysArgGlyMetThrLysAspHisThrPAspHis 627
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1695 GGGAGGGCTAATGAAGCGAGCATGTCAAAGACTTTCACGTGGGACCAT 1744

628 Ala 628
|||
1745 GCC 1747

seq_name: /SIDSl/gcgdata/geneseq/geneseq/NA1997.DAT: AAT67285

seq_documentation_block:
ID AAT67285 standard; cDNA; 2992 BP.
XX
XX AAT67285;
XX
DT 11-SEP-1997 (first entry)
XX
DE Soluble starch synthase cDNA clone SSS10.52.
XX
XX Soluble starch synthase; SSS10.52; transgenic plant; cereal; maize;
KW ss.
XX
OS Zea mays inbred line W64A.
XX
XX W097209336-Al.
XX
XX 12-JUN-1997.
XX
XX 04-DEC-1996; 96W0-GB02990.
XX
XX 06-DEC-1995; 95GB-0024938.
XX
XX (ZENE ) ZENECA LTD.
XX
XX Keeling PL, Knight ME;
XX
XX WPI; 1997-319782/29.
XX
XX cDNA encoding soluble starch synthase - used to produce transgenic
XX plants with increased capacity for producing and storing starch
XX
XX Claim 5; Page 17-20; 44pp; English.
XX
XX cDNA clones SSS10.52, SSS6.31 and SSS56 (AAT67285-87) code for maize
XX soluble starch synthase (SSS). They were isolated from a maize
XX inbred line W64A library by screening with a probe (AAT67288) based
XX on pea SSS, rice SSS and maize GBSS. The cDNAs can be used to
XX produce plants with an increased capacity for producing starch, or
XX a capacity to produce starch with an altered fine structure. They
XX can also be used to isolate the corresponding genomic sequences
XX from crop plants, to determine the contribution of the SSS gene to
XX the net regulation of starch biosynthesis, and to modify the levels
XX of starch produced by the plant. Transgenic plants (esp. maize)
XX can be used to produce hybrid plants which have higher rates of
XX starch synthesis at temperatures above the normal optimum.
XX
XX Sequence 2992 BP; 758 A; 655 C; 801 G; 776 T; 2 other;
XX

```



```
287 AlaserLeuValProValLeuAlaAlaLysTyrArgProTyrGlyVa 303
1440 GCCAGTCTAGACCCAGTCTCTTCTGCTGCAAAATATAGCCATATGGTGT 1489
303 lTyrArgaspSerArgSerThrLeuValIleHisAsnLeuAlaHisGlnG 320
1490 TTATAAGACATCCCGCAGCATCTTGTATAACATAAATTAGCACATCAGG 1539
320 lYValGluProAlaSerThrTyrProaspLeuGlyLeuProProGluTrp 336
1540 GTGTAGACCCTCGACAGCACATATCTGACCTTGGGTGGCCACCTGAATGG 1589
337 TyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAlaLe 353
1590 TATGGAGCTCTGGAGTGGGTATTCCTGCAATGGCGAGGAGCATGCCCT 1639
353 uAspLysGlyGluAlaValasnPheLeuLysGlyAlaValValThrAlaA 370
1640 TGACAAGGGTGAGGCAGGTAAATTTTGAAGAGGTGCAGTTGTGACAGCAG 1689
370 sPArgIleValThrValSerGlnGlyTyrSerTrpGluValThrThrAla 386
1690 ATCGAATCGTCAGTGTCAAGGGTATTTCATGGGAGGTCCAACTGCT 1739
387 GluGlyGlnGlyLeuAsnGluLeuSerSerArgTysSerValLe 403
1740 GAAGGTGACAGGGCCTCAATGAGTCTTAAAGTCCAGAAAGAGTGATT 1789
403 uAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrThra 420
1790 AAACGGAATGTAATGAATGACATTAAATGATTGGAACCTGCCACAG 1839
420 sPLysCysLeuProHisHisTyrSerValaspAspLeuSerGlyLysAla 436
1840 ACAAAATGATATCCCTGTCTATTCTGTGTGATGACCTCTCTTGAAGGCT 1889
437 LysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAs 453
1890 AAATGTAAGGTGATGTCGACGAAGGAGCTGGGTTTACCTATGAAGGCTGA 1939
453 pValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleA 470
1940 TGTTCCTCTGATGGCTTTATTGGAAGATTGGATTATCAGAAAGGCATTG 1989
470 sPLeuIleLysMetalaleProGluLeuMetArgGluAspValGlnPhe 486
1990 ATCTCATTTCAACTTATCATACCATCTCATCGCGAAGAATGTTCAATTT 2039
487 ValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgSerTh 503
2040 GTCATGCTTGATCTGTGACCCAGAGCTTGAAGATTGGATGAGATCTAC 2089
503 rGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerValp 520
2090 AGAGTCGATCTTCAAGGATAAATTCGTGGATGGGTGGATTAGTGTTC 2139
520 roValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSer 536
2140 CAGTTTCCCAACCGAATAACTCGCGCTGGCGATATATTGTTAATGCCATCC 2189
537 ArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyTh 553
2190 AGATTGCAACCTTGTGTGTCATACATGAGCTATATGCTATGAGATGGCAC 2239
553 rValProValValHisGlyThrGlyGlyLeuArgAspThrValGluThrP 570
2240 AGTTCCTGTTGTCATGCAACTGGGGCCTTAGAGATACCGTGGAGAACT 2289
570 heAsnProPheGlyAlaLysGlyGluGlyThrGlyThrGlyTrpAlaPheSer 586
2290 TCAACCCCTTTCGGTGAAGATGGAGAGCAGGCTACAGGGTGGGCATTCGCA 2339
587 ProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSerT 603
```

```
2340 CCCCTAACCCAGAGAAACATGTTTGTGACATTCGAACTGCAATATCTA 2389
603 hrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMet 619
2390 CATACAGGGAACACAAAGTAATAATGGGAAGGCTAATGAAGCCAGGCATG 2439
620 ThrLysAspHisThrTrpAspHisAlaProSerSerThrSerArgSers 636
2440 TCAAAAGAGTTTCAGTGGGACCATGCCCTGAAACAATACGAACAATCTT 2489
636 erSerGlyProSerTrpThrAsnProThrSerCysArgArgGlyLeuG1 652
2490 CCATGGCCCTTCATCGGATCGACCC...GATGTTCAATGAAAAAGGG 2536
652 yArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSerSerp 669
2537 ACCAAAGTGTGTTGCTTCCCTTGAAGATTATCAGTTCATCATCTTATAGTA 2586
669 heArgGlyProGluGlyTyrPro...CysThrLeuArgCysProAla 683
2587 AGCTGAATGATGAAGAAACCCCTGTACATTACATGGAAGGCAGACCGG 2636
684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPhe... 695
2637 CTATTGGCTCCATTGCTCCCAATGCTGCTTGGCTGCTTCCCTCCCTCGATGG 2686
695 ..... 695
2687 ACCGGATGTCAGTGAGGAATCCAGNCGACAGACAGTTCCTTGAAGGATAGGAA 2736
696 .....AlaGlySerArgThrTyrAspGlyCysAlaAlaAlaValThr 710
2737 GGGAGCTGGAAGC...AGTCACGCGAGGCGAGGCAAGCCTTCGCGCTTAAT 2783
711 AlaSerGlyGlyArgGlnLeuGlnPhe 719
2784 TCATATGGAACAAGCTGGAGTCAGTTT 2810
seq_name: /SIDS1/gc9data/geneseq/geneseq/NA1998.DAT.AAV66832
seq_documentation_block:
ID AAV66832 standard; DNA; 2990 BP.
XX
AC AAV66832;
XX
DT 05-JAN-1999 (first entry)
XX
DE Zea mays soluble starch synthase gene SSSI052 and SSS64.
XX
KW Zea mays; US yellow-dent corn line; maize; soluble starch synthase;
KW SSS; glycogen biosynthetic pathway; branching enzyme; ss.
XX
OS Zea mays.
XX
PN US5824790-A.
XX
PD 20-OCT-1998.
XX
PF 15-DEC-1995; 950S-0572951.
XX
PR 15-DEC-1995; 950S-0572951.
PR 21-JUN-1994; 94US-0263921.
PR 29-NOV-1994; 94US-0346602.
XX
PA (ZENE ) ZENECA LTD.
XX
PI Guan H, Keeling PL, Knight ME;
XX
DR WPI; 1998-582626/49.
XX
PT Isolated nucleic acid molecule, used to produce transgenic plants -
comprises nucleotide sequence encoding polypeptide having soluble
```



starch synthase activity, where polypeptide is encoded by maize gene  
 Claim 1; Column 25-28; 29pp; English.  
 The present sequence represents an isolated nucleic acid molecule which  
 has been isolated and comprises a nucleotide sequence encoding a  
 polypeptide having soluble starch synthase (SSS) activity, where the  
 polypeptide is encoded by a maize gene. The isolated nucleic acid  
 molecule can be used to produce transgenic plants with altered starch  
 production. The transgenic plants produced using the nucleic acid  
 molecule have an enhanced ability to produce structurally-altered starch.  
 Sequence 2990 BP; 758 A; 655 C; 801 G; 776 T; 0 other;

```

alignment_scores:
  Quality: 2651.00      Length: 643
  Ratio: 4.587          Gaps: 7
  Percent Similarity: 89.891  Percent Identity: 79.471

alignment_block:
  US-09-674-824-2 x AAV65832  ..

Align seg 1/1 to: AAV66832 from: 1 to: 2990

6 ValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAspProAl 22
  ::||| ||| ::||| ||| |||
613 CTCGGCGGTGGCGCGCGGTGCTCTCTCGCGGGCGCCTGGCGCGC 662
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
22 aThrAlaAla...ArgAlaSerAlaCysValValArgAlaArgLeuArgA 38
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
663 CGCGGTGGCGGACGCGCGCGCGCGG.....CGAGGCTCCAGC 700
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
38 rglLeuAlaArgGlyArgTyrValAlaGluLeuSerArgGluGlyProAla 54
  ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
701 GCGTGTCTGCGCGCGCGCGGTGCGTTCGGAGCTGAGCAGGAGGGCGCGC 750
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
55 AlaArgProAlaGlnGlnGlnLeuAlaProProLeuValProGlyPh 71
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
751 CGCGCGCGCGCTGCCACCGCGGTGCTGGCGCGCGCGCTCTGTCGCGC 800
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
71 eLeuAlaProProProAlaProAlaGlnSerProAlaProThrGlnP 88
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
801 CCTCGCG...CGCGCGCGCGAGCGACGGGTGAGCGCGCATCGACCCG 847
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
88 rProLeuProAspAlaGlyValGlyGluLeuAlaProAspLeuLeu 104
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
848 CGCGCGTGGCGACGCGCGCTGGGGACCTCGGT.....CTCGAACCT 891
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
105 GluGlyIleAlaGluAspSerIleAspSerIleIleValAlaAlaSer 121
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
892 GAAGGGATTGCTCAAGGTTCATCGATACACAGTAGTAGTTGGCAAGTGA 941
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
121 uGlnAspSerGluIleMetAspAlaAsnGluGlnProGlnAlaLysVal 138
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
942 GCAAGATTCTGAGATTGTGTGGGAAGGAGCAAGCTCGAGCTAAAGTAA 991
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
138 hrArgSerIleValPheValThrGlyGluAlaAlaPro.TyrAlaLys 154
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
992 CACAAGCATGTCTTTGTAAACCGCGGAGCTTCTCCTTAATCGAAGTC 1041
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
154 rGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuAlaAla 171
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
1042 TGGGGGTCTAGGAGATGTTGTGGTTCATTGCCAGATTGCTCTTGTGCT 1091
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
171 rGlyIleHisArgValMetValValMetProArgTyrLeuAsnGlySer 187
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
1092 GCGGTCAACCGTGTGATGTTGTTAATGCCACAGCATTTAAATGGTACC 1141
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
188 AspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIle 204
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
1142 GATAAGAAATTATGCAATGATATTTTACTCAGAAGAACACATTCGGATT 1191
  ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||

```

204 cCysPheGlyGlySerHisGluValThrPhePheHisGluTyrArgAspA 221  
1192 ATTCCTTTGGCGGTGAACATGAAGTTACCTCTCTCCATGAGTATAGAGATT 1241  
221 snValAspTrpValPheValAspHisProSerTyrHisArgProGlySer 237  
1242 CAGTTGACTGGGTGTTGTTGATCATCCTCATATACAGACCTGGAAAT 1291  
238 LeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrTh 254  
1292 TTATATGAGATAAGTTGGTGCCTTTGGTGATAAATCAGTTTCAGATACAC 1341  
254 rIeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyG 271  
1342 ACTCCTTGCATGCTGCATGTGAGGTCTCTTGGCTTGAATGGGAG 1391  
271 lYrTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrpHisAla 287  
1392 GATATATTATGGACAGAATTGCATGTTGGTTGTCAATGATTGGCATGCC 1441  
288 SerLeuValProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTy 304  
1442 AGCTCTAGAGCCAGCTCTTCTTCTGCGCAAAATATAGACCATATGGTGTTA 1491  
304 rArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyV 321  
1492 TAAAGAGCTCCCGCAGCATCTCTGTAAATACATAATTAGCACATCAGGGTG 1541  
321 aIGluProAlaSerThrTyrProAspLeuGlyLeuProGluTyrpTyr 337  
1542 TAGAGCCTGCAAGCACATATCCTGACCTTGGGTGGCCACCTGAATGGTAT 1591  
338 GlyAlaLeuGluTyrpValPheProGluTrpAlaArgArgHisAlaLeuAs 354  
1592 GGAGCTCTGGAGTGGGTATTCCTCGTAATGGCGAGGAGGCATGCCCTTGA 1641  
354 pLysGlyGluAlaValAsnPheLeuLysGlyAlaValValThrAlaAspA 371  
1642 CAAGGCTGAGGCAGCTAATTTTTGAAAGGTGCAGTTGTCACAGCAGATC 1691  
371 rGlyLeuValThrValSerGlnGlyTyrSerTrpGluValThrThrAlaGlu 387  
1692 GAATCGTGACTGTCAAGTAAAGGGTTATTCTGCGGAGGTCAACAACCTGCTGAA 1741  
388 GlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAs 404  
1742 GGTGACAGGGCCCTCAATGAGCTCTTAAGCTCCAGAAAGAGTGTATTAAA 1791  
404 nGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspL 421  
1792 CGGAATTGTAAATGGAATTGCATTAATGATGTGAACCCCTGCCACAGACA 1841  
421 ySCysLeuProHisHisTyrSerValAspAspLeuSerGlyLysAlaLys 437  
1842 AATGATATCCCTGTCATTAATTCTGTTGATGACCTCTCTTGAAGGCTAAA 1891  
438 CysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAspVa 454  
1892 TGTAAAGGTGCATTCGCAAGGAGGTGGTATTACTATAAGCCCTGATGT 1941  
454 lProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspL 471  
1942 TCCTCTGATTTGGCTTATTGGAAAGATTGGATTATCAGAAGGCATTGATC 1991  
471 euIleLysMetAlaIleProGluLeuMetArgGluAspValGlnPheVal 487  
1992 TCATCAACATTTATCATACCATCTCATCGCGAAGAATGTTCAATTTGTC 2041  
488 MetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgSerThrGl 504  
2042 ATGCTTTGGATCTGGTGACCCAGAGCTCGAAGATTGGATGGAGATCTACAGA 2091  
504 uSerSerTyrLysAspLysPheArgGlyTrpValIcLysPheSerValProV 521







|||||  
554 GACAGAATTGCAATGTTTGTGTTCAATGATTGGCATGCCAGCTAGTGCCA 603  
292 ValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSerAr 308  
|||||  
604 GTCCTCTCTGCTGCAAAATATAGACCATATGGTGTATTATAAGACATCCCG 653  
308 gSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProAlas 325  
|||||  
654 CAGCATCTCTGTAATACATAATTTAGCACATCAGGGGTGTAGAGCCTGCAA 703  
325 erThrTyrProAspLeuGlyLeuProGluTyrProGluTyrGlyAlaLeuGlu 341  
704 GCACATATCCTGACCTTGGGTGGCAGCTGAATGATGGAGCTCTGGAG 753  
342 TrpValPheProGluTyrAlaAlaArgHisAlaLeuAspLysGlyGluAl 358  
754 TGGGTATTCCCTGAATGGCGCAGGAGCATGCCCTTGACAAGGGTGAAGC 803  
358 aValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValThrV 375  
804 AGTTAAATTTTGAAGGTGCGAGTGTGCACACGACAGATCGAATCGTGACTG 853  
375 alSerGlnGlyTyrSerTrpGluValThrAlaGluGlyGlyGlnGly 391  
854 TCAGTAGGGTTATTCGTGGGAGGTGCACACTGCTGAAGGTGGACAGGC 903  
392 LeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleValAs 408  
904 CTCATAGAGCTCTTAAGCTCCAGAAGAGTGTATTAAACGGAATTTGTA 953  
408 nGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuProH 425  
954 TGAATTTGACATTAATGATTGAACCTCTGCGACAGCAAAATGTATCCCT 1003  
425 lshIstYrSerValAspAspLeuSerGlyLysAlaLysCysLysAlaGlu 441  
1004 GTCAATTATTCTTGTGATGACCTCTCTGGAAGGCCAAATGTAAGGTGCA 1053  
442 LeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuIleGl 458  
1054 TTGCAGAGAGGAGCTGGTTTACATATAGGCTGATGTTCTCTCTGATGG 1103  
458 yPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMeta 475  
1104 CTTTATTGGAAGTTGGATTATCAGAAAGGCATTGATCTCATTTCAACTTA 1153  
475 laileProGluLeuMetArgGluAspValGlnPheValMetLeuGlySer 491  
1154 TCATACAGATCTCATGCGGGAGATGTTCAATTTGTCATGCTTGGATCT 1203  
492 GlyAspProIlePheGluGlyTrpMetArgSerThrGluSerTyrIly 508  
1204 GGTGACCCAGAGCTTGAAGATGGATGAGATCTACAGATCGATCTTCAA 1253  
508 sAspLysPheArgGlyTrpValGlyPheSerValProValSerHisArgI 525  
1254 GGATAAATTTCTGGATGGTGGATTAGTTAGTCTCCAGTTTCCACCGAA 1303  
525 leThrIleGlyCysAspIleLeuLeuMetProSerArgPheGluProCys 541  
1304 TAACTCCGCGCTCGCATATATTGTTAATGCCATCCAGATTCGAACCTCT 1353  
542 GlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValIleH 558  
1354 GGTCTCAATCAGCTATATGCTATGATGATGACAGTTCCTGTTGTCCA 1403  
558 sGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheGlyA 575  
1404 TGCAACTGGGGGCTTAGAGATACCGTGGAGAAGTCAACCCCTTTGCGTG 1453  
575 laLysGlyGluGluGlyThrGlyThrAlaPheSerProLeuThrValAsp 591  
|||||

1454 AGAATGGAGCAGCGGTACAGGTGGGCATTCGCACCCCTCAACACAGAA 1503  
592 LysMet. LeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisL 608  
|||||  
1504 AACATGTTTGTGGACATTCGGAACCTGCAATATCTACATACAGGGAACA 1553  
608 ysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisThr 624  
1554 AGTCTCTCTGGGNAAGGCTAATGAACGAGCATGTCAAAAGACTTTCAG 1603  
625 TrpAspHisAla 628  
1604 TGGGACCATGCC 1615  
seq\_name: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ50637  
seq\_documentation\_block:  
ID\_AAZ50637 standard; cDNA; 1528 BP.  
XX  
AC\_AAZ50637;  
XX  
DT 23-MAY-2000 (first entry)  
XX  
DE Corn soluble starch synthase gene fragment inserted in pSS42.  
XX  
KW Soluble starch synthase; starch fine structure; corn;  
KW transgenic plant; amylose; amylopectin; amylose polymerisation;  
KW non-granule bound starch synthase; non-GBSSI; altered starch; food;  
KW paper; plastic; adhesive; ss.  
XX  
OS Zea .mays.  
XX  
PN WO200006755-A2.  
XX  
PD 10-FEB-2000.  
XX  
PF 26-JUL-1999; 99WO-US16296.  
XX  
PR 28-JUL-1998; 98US-0094436.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
PI Broglie KE, Lightner JE;  
XX  
DR WPI; 2000-195311/17.  
XX  
PT Producing transgenic cereal crops with altered starch structure useful  
PT for preparing foodstuff, paper, plastic or adhesives, comprises  
PT transforming crops with chimeric sense or antisense gene construct  
PT encoding starch synthase  
XX  
PS Claim 5; Page 51-52; 56pp; English.  
XX  
CC The present sequence is the corn soluble starch synthase (SSI) DNA  
CC sequence inserted into plasmid pSS42. Chimeric gene of pSS42 was cloned  
CC into the vector pKS17 for the generation of an antisense construct for  
CC suppression of SSI expression in corn. The starch fine structure derived  
CC from a grain of the cereal crop can be altered in the transformed cereal  
CC crop by changes in amylose to amylopectin ratio, amylopectin fine  
CC structure, increased abundance of very short amylopectin chains and in  
CC the degree of polymerisation of amylose. These modifications can be  
CC created by controlling the expression of non-GBSSI (non-granule bound  
CC starch synthase) in transgenic plants. Altered starches are useful in  
CC foods, paper, plastics or adhesives.  
XX  
SQ Sequence 1528 BP; 440 A; 383 C; 300 G; 405 T; 0 other;

alignment\_scores:  
Quality: 2495.00 Length: 508  
Ratio: 5.051 Gaps: 1  
Percent Similarity: 97.244 Percent Identity: 89.567



alignment\_block:

US-09-674-824-2 x AA250637/rev ..

Align seg 1/1 to reverse of: AA250637 from: 1 to: 1528

```

134 GlnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPr 150
1527 CGAGCTAAAGTAAACAACAAGCATTTGCTTTGTAACCGGCGAAGCTTC 1478
150 oTyAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleA 167
1477 TTATGCAAACTCTGGGGGCTAGGAGATGTTTGTGTTCAATGGCCAGTTG 1428
167 laLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeu 183
1427 CTCCTTGCTGCTGCTGCTCACCGTGTGATGTTGTAATGCCAGATATTTA 1378
184 AsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHI 200
1377 AATGGTACTCCGATAAGAAATTTATGCAATGCATTTTACAGAAAAACA 1328
200 sIleLysIleProCysPheGlyGlySerHisGluValThrPhePheHISG 217
1327 CATTCGGATTCCATGCTTTGGCGGTGAACATGAAGTTTACCTTCTTCCATG 1278
217 luTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHis 233
1277 AGTATAGAGATTACGTTGACTGGGTGTTTGTGATCATCCCTCATATCAC 1228
234 ArgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGI 250
1227 AGACCTGGAAATTTATATGGAGATAAGTTTGGTCTTTGTGATAATCA 1178
250 nPheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleL 267
1177 GTTCAGATACACACTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1128
267 euGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsn 283
1127 TTGAATGGGAGGATATATTTATGGACAGAAATTCATGTTTGTGTCAAT 1078
284 AspTrpHisAlaSerLeuValProValLeuLeuAlaLysTyrArgPr 300
1077 GATTGGCATGCGACTCTAGTGCCAGTCTTCTTCTGCTGCAAAATATAGACC 1028
300 oTyGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuA 317
1027 ATATGGTGTATTATAAGACTCCCGCAGCATTCCTTGTATAACATAATTAG 978
317 laHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuPro 333
977 CACATCAGGTGTAGAGCTGCAAGCACATATCTTGACCTTGGGTGCCA 928
334 ProGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgAr 350
927 CCTGAATGGTATGGAGCTCTGGAGTGGTATTCCCTGAATGGCGGAGGAG 878
350 qHisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValV 367
877 GCATGCCCTTTGACAAAGGCTGAGGAGTAAATTTTGAAGGTGCAGTTG 828
367 alThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluVal 383
827 TGACAGCAGATCGAATCGTCACTGCTCAGTAAGGGTTATTTCGTGGGAGGC 778
384 ThrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLy 400
777 ACAACTGCTGAAGGTGACAGGCGCTCAATGAGCTCTTAAGCTCCAGAAA 728
400 sSerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnP 417
727 GAGTGTATTAAACGGAATGTAAATGGAATTGACATTAATGATTGGAACC 678

```

```

417 roThrThrAspLysCysLeuProHisHisTyrSerValAspLeuSer 433
677 CTGCCACAGACAAATGTATCCCTGTCTATTATTCTGTTGATGACCTCTCT 628
434 GlyLysAlaLysCysLysAlaLeuGluGlnLysGluLeuGlyLeuProVa 450
627 GGAAGGCCAAATGTAAAGGTGCATTGCAGAAAGGAGCTGGGTTTACCTAT 578
450 lArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnL 467
577 AAGCCTGATGTTCTCTGATGCTGCTTATTGGAAGGTGCGATTATCAGA 528
467 ySGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAsp 483
527 AAGCATGTGATCTCAATCAACTTATCATACCAGATCTCATCGCGGAAGAT 478
484 ValGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMe 500
477 GTTCAATTTGTCATGCTTGGATCTGGTGAGCCAGAGCTTTGAAGATTGGAT 428
500 tArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyP 517
427 GAGATCTACAGAGTCGATCTTCAAGGATAAATTTCTGGTGGATGGTTGGAT 378
517 heSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeu 533
377 TTAGTGTTCAGTTTCCACCAGATAACTCCCGCTGCGATATATTGTTA 328
534 MetProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGI 550
327 ATGCCATCCAGATTCGAACTTGTGCTCAACATCAGCTATATGCTATGCA 278
550 nTyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrV 567
277 GTATGGCACAGTTCTCTGTTCTCCATGCAACTGGGGGCTTAGAGATACCG 228
567 alGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrp 583
227 TGAGAGAACTTCAACCTTTTCGGTGAGATGGAGAGGAGTACAGGGTGG 178
584 AlaPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAl 600
177 GCATTGCGACCCCTAACACACAGAAAACATGTTGTGGACATTGCGAACTGC 128
600 aMetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysA 617
127 AATATCTACATACAGGGAACACAAAGTCTCTCTGGGAAGGCTAATGAAGC 78
617 rgGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSe 633
77 GAGGCATGTCAAAAGACTTCACGTGGGACCATGCCCTGAAACAATACGAA 28
633 rArgSerSerSerGlyProSer 640
27 CAAATCTTCCAGTGGGCCCTTCA 6

```

seq\_name: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA34652

seq\_documentation\_block:

ID AA34652 standard; DNA; 10336 BP.

XX AA34652;

XX DT 05-JUL-1999 (first entry)

XX DE Wheat starch soluble synthase I (SSS I) gene sequence.

XX KW Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBSS;  
 KW starch branching enzyme; starch soluble synthase; debranching enzyme;  
 KW endosperm; wheat; barley; granule-bound synthase; glutenin; starch;  
 KW grain softness protein I; bacterial isoamylase; glycogen synthase;  
 KW WSBF I-D4 gene; ds.



OS Triticum tauschii.  
 XX Key Location/Qualifiers  
 FH CDS 1..9713  
 FT /tag= a  
 FT /product= "SSS I"  
 FT /note= "Contains introns"  
 FT 1..316  
 FT /tag= b  
 FT /number= 1  
 FT 317..1471  
 FT /tag= c  
 FT /number= 1  
 FT 1472..1828  
 FT /tag= d  
 FT /number= 2  
 FT 1829..2765  
 FT /tag= e  
 FT /number= 2  
 FT 2766..2823  
 FT /tag= f  
 FT /number= 3  
 FT 2824..2905  
 FT /tag= g  
 FT /number= 3  
 FT 2906..3028  
 FT /tag= h  
 FT /number= 4  
 FT 3029..4112  
 FT /tag= i  
 FT /number= 4  
 FT 4113..4194  
 FT /tag= j  
 FT /number= 5  
 FT 4195..4285  
 FT /tag= k  
 FT /number= 5  
 FT 4286..4459  
 FT /tag= l  
 FT /number= 6  
 FT 4460..4561  
 FT /tag= m  
 FT /number= 6  
 FT 4562..4643  
 FT /tag= n  
 FT /number= 7  
 FT 4644..4743  
 FT /tag= o  
 FT /number= 7  
 FT 4744..4855  
 FT /tag= p  
 FT /number= 8  
 FT 4856..4998  
 FT /tag= q  
 FT /number= 8  
 FT 4999..5021  
 FT /tag= r  
 FT /number= 9  
 FT 5022..5101  
 FT /tag= s  
 FT /number= 9  
 FT 5102..5192  
 FT /tag= t  
 FT /number= 10  
 FT 5193..8592  
 FT /tag= u  
 FT /number= 10  
 FT 8593..8718  
 FT /tag= v  
 FT /number= 11  
 FT 8719..8806  
 FT /tag= w  
 FT /number= 11

FT exon 8807..8915  
 FT /tag= x  
 FT /number= 12  
 FT 8916..8991  
 FT /tag= y  
 FT /number= 12  
 FT 8992..9104  
 FT /tag= z  
 FT /number= 13  
 FT 9105..9160  
 FT /tag= aa  
 FT /number= 13  
 FT 9161..9199  
 FT /tag= ab  
 FT /number= 14  
 FT 9200..9497  
 FT /tag= ac  
 FT /number= 14  
 FT 9498..9713  
 FT /tag= ad  
 FT /number= 15  
 XX  
 PN WO9914314-Al.  
 XX  
 PD 25-MAR-1999.  
 XX  
 PF 11-SEP-1998; 98WO-AU00743.  
 XX  
 PR 20-MAR-1998; 98AU-0002509.  
 PR 12-SEP-1997; 97AU-0009108.  
 XX  
 PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
 PA (GOOD-) GOODMAN FIELDER LTD.  
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.  
 PA (AUSU ) UNIV AUSTRALIAN NAT.  
 XX  
 PI Li Z, Morell M, Rahman S;  
 XX  
 DR WPI; 1999-229525/19.  
 DR P-PSDB; AAY09004.  
 XX  
 PT New isolated cereal plant enzyme genes used for, e.g. expression of  
 PT antisense sequences of granule bound synthase  
 XX  
 PS Claim 10; Page 87-94; 171pp; English.  
 XX  
 CC The invention relates to a novel enzyme of starch biosynthetic pathway  
 CC in a cereal plant, where the enzyme is selected from starch branching  
 CC enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching  
 CC enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or  
 CC SBE I of rice or maize. The methods and products can be used for  
 CC targeting expression specifically to the endosperm of the seeds of cereal  
 CC plants such as wheat or barley. They can be used for the expression of  
 CC e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low  
 CC mol. wt. glutenin, grain softness protein I, bacterial isoamylase,  
 CC bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They  
 CC can be used for modifying the characteristics of starch produced by a  
 CC plant. The present sequence represents the wheat SSS I gene sequence.  
 XX  
 SQ Sequence 10336 BP; 2733 A; 2055 C; 2625 G; 2921 T; 2 other;

alignment\_scores:  
 Quality: 2349.00 Length: 3317  
 Ratio: 3.149 Gaps: 25  
 Percent Similarity: 22.490 Percent Identity: 22.339  
 alignment\_block:  
 US-09-674-824-2 x AAX34652 ..  
 Align seg 1/1 to: AAX34652 from: 1 to: 10336  
 1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17



```

|||||
1  ATGGCGCGCAGCGGCTCGCGCGCGGTCGCTCGCCCGCAGCGCTCGGCT 50
17  uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
51  GCGCGCGGATCCGGCGAGCGCGCGCGCGCGGCTCGCGCTCGCGTCCGCG 100
34  laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
101  CGCGGCTCGGCGCTTGGCGGGGCGCGCTACGTCGCCGAGCTCAGCAGG 150
51  GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnLeuAlaProProLe 67
151  GAGGCGCGCGCGCGCGCGCGCGCGCGCAGCAGCAACTGGCCCGCGCGCT 200
67  uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84
201  CGTGCCAGCGCTTCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 250
84  laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
251  CCCCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
101  AspLeuLeuLeu Glu 105
301  GACCTCCTGCTCGAAGGTAAAAACAAGGCTGAATCCTCAGATCACCTCG 350
105  ..... 105
351  CGCTCTCGTTTTACAAATACGGTACTCGGNAAGTGGTGGCTGTATATGA 400
105  ..... 105
401  AGTTCTGTCGATTCTTCCTGACGGATGTTTCAGTCGATTTCAGTTGTATA 450
105  ..... 105
451  TATGTATACGTTCTGTTTCATCGATCGTACAGATTTCACGACACTA 500
105  ..... 105
501  GATAGAAATCGAGACCGACGCGGCGAGATCAATAGATTTTCTAGAGGTT 550
105  ..... 105
551  TTATTGGATCGTAGATGATTGATGGGGTGGCGGTGTCGATACGATAGCG 600
105  ..... 105
601  GTGCACCGCGGATGATATCGGGGATGTGCACGTGGTGGGTCTCAGCAGA 650
105  ..... 105
651  CATATCACTAGACTGGTATCGTAATTACTAGTACTACTGGAAGAGGAC 700
105  ..... 105
701  TAAAAAGGCTAGGCAAGTGCACGATGTTGGCAACGTTCTTAAATGTAT 750
105  ..... 105
751  GAGTTTGTCTTTGCTGGGCTGGTATTATTACCAAAAAAATGGTGTAGT 800
105  ..... 105
801  CCCTGTACTTATTAAATGGGNAATCTTAACATGACACTGGGGTTTATGAG 850
105  ..... 105
851  TCTCCAATTGTATATCTCAGCACTCAACTGATTTTACTGATACGTACT 900
105  ..... 105

```

```

901  GAAATGACACGTGAGCACCCCTTCAAGGAATGCAATGCTTCTTCTG 950
105  ..... 105
951  TTTTATATTACAGGAACCTAGAGGAGCTTCCACCTTTGAGTACAGAA 1000
105  ..... 105
1001  CTCCTCCGTTCCAAAAATAGATGACTCAACTTGTCTACTAATTTGTACTA 1050
105  ..... 105
1051  TAGTTAGTACAAAGTTGAGTCATCTATTTAGAACGGAGGAGTAGTATC 1100
105  ..... 105
1101  GAAATTGAAGACCTTGTATTACTGTCTTGTCTTTCATCAATAAATGGGAG 1150
105  ..... 105
1151  GCCCATGCAGTAAGTCACATGGGCACCTGGGAGGCTGGGATCATGTGTGC 1200
105  ..... 105
1201  TTTGCAGAGTACTAGACCCAGCTCACCTCTGTTAGATTACTTGTGGGC 1250
105  ..... 105
1251  TGCTACTTTGTGTTGCTGTGCAGTATATCAGACATCCTGAATTTGGCAT 1300
105  ..... 105
1301  CTAGCTGAGAACAGAAATGCGAGGTGCGCCCATCTTATTATTGCTAAACTG 1350
105  ..... 105
1351  TTGTCACGCAATTTATAAAGAAATGTGATCTTCTGAGTATTAAATTAATCAT 1400
105  ..... 105
1401  GTTCTGCTAATATCTCTCTCTCTCTGTTGGTGGTGCACAAATATACCATATGA 1450
106  ..... GlyIleAlaGluAspSerIleAspSerIle 115
1451  ATATTTTCCATTTTGCAACACAGGATGCTGAGGATTCATCTCAGACAGCAT 1500
115  eIleValAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluG 132
1501  AATCGTGGCTGCAAGTGACGAGGATTCAGATCATGGATGCAATGAGC 1550
132  InProGlnAlaLysValThrArgSerIleValPheValThrGlyGluAla 148
1551  AACCTCAAGCTAAAGTTACACGTAGCATCGTGTGTTGACTGGTGAAGCT 1600
149  AlaProTyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuPr 165
1601  GTCCTTATGCAAGTCAAGGGGCTGGGAGATGTTGTGTGTTCTGTTACC 1650
165  oIleAlaLeuAlaAlaArgGlyHisArgValMetValValMetProArgT 182
1651  AATTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700
182  yLeuAsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAla 198
1701  ACTTGAATGGGTCTCTGTATAAAACTATGCAAGGCATTATACACTCGG 1750
199  LysHisIleLysIleProCysPheGlyGlySerHisGluValThrPhePh 215
1751  AAGCATATTAGATTCCATGCTTTGGGGATCAGATGAAGTGACCTTTT 1800
215  eHisGlyTyrArgAspAsnValAspTrp..... 224
1801  TCATGAGTATAGAGACAACGTCGATTGGTGGGTACACAATCACCTTCCT 1850

```



224 ..... 224  
1851 ATTCTGTGTAATGTAGCAACTGTTTATCCCTGTTTACACTTCCTTTA 1900  
224 ..... 224  
1901 GCCCTGCAAGACATATGTGATTTCCATACITTTTTTGTATTATCCCTGT 1950  
224 ..... 224  
1951 ACTCTTGCTCATGAAGTCAAAATATCATATATCATCCATGAAGTCATGCAT 2000  
224 ..... 224  
2001 GTGCCTAGTATTTTTGGTGTGCGGTGCCCTTAACTTTTCAGGGATTAAATACG 2050  
224 ..... 224  
2051 TGGAAATTTGATAACTAAAGTTTTATTTTATGAAAAAAATTCAGTTGGT 2100  
224 ..... 224  
2101 GAGCCACAGCCACGACGTGSCACCACCTGCTTGACATGATTTTGCATTT 2150  
224 ..... 224  
2151 CTGTTTGACCGGACGACTTCATGTGAATAAGGTGTAAATCATAAAGTAC 2200  
224 ..... 224  
2201 CAATTTTATCTGCCAATTCGACTTAAGAGTATATACATTTATCTTGGCC 2250  
224 ..... 224  
2251 TCAATCATGGGAGTACTGTGCATTTCAGTGCACCATCATTTGTTCTAAGGAG 2300  
224 ..... 224  
2301 AAAATGTGGTGCAAGGAAGACACTTTTGTCCCTTAATAAAGGCGAGCA 2350  
224 ..... 224  
2351 CTCCTGTGTGCATATAGATAGAAGCAACAACCTTATTTCAAAGAGCTAAC 2400  
224 ..... 224  
2401 AATGCCAAAGAACCMAAANAAGCATGCTAAGCGGTGACACCAAAAGGT 2450  
224 ..... 224  
2451 GAGGGGGCCTTGTGACTGACAGCACCCCAAACTATTGCCATTGTTTAC 2500  
224 ..... 224  
2501 TAAATGAAGATCATTTTAGAAGCTCTCAGGAACCTTCGAAAACAGTGGCTT 2550  
224 ..... 224  
2551 TCCGTCACAGATCGTCTGTTAATATTTTGTCCAGTGATACTTTTTTTG 2600  
224 ..... 224  
2601 CTCTTTACAAGAGTGCCTATGTTGACATATACATTTGTTAAGTTGTTTACA 2650  
224 ..... 224  
2651 AGTTTACTTCTTATCTAAACAGCAAGTGCCTAAATGCTTGCATTATTTT 2700  
225 ..... Val. Phe 226  
|||  
2701 GGCTATTTATTTTATCTCATTTCAATCAACACTTTTGTTCAGGTGTTT 2750

227 ValAspHisProSerTyrHisArgProGlySerLeuTyrGlyAspAsnPh 243  
|||  
2751 GTCGATCATCCGTCAATATCATAGACAGGAAGTTTATATGGAGATAATTT 2800  
243 eGlyAlaPheGlyAspAsnGln. Phe. .... 251  
|||  
2801 TGGTGTCTTTGGTGATTAATCAGGTACACTACACTATACTAAGCTCCTAGT 2850  
251 ..... 251  
2851 TGACTAAGTCGTAAGTTGTACCTCCCTCGCTGACCGGCTGCTCTATGTCGT 2900  
252 ..... ArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuI 266  
|||  
2901 GCAGTTTCAGATACACACTCCTTTTGTATGCTGCTGCGAGGCCCACTAA 2950  
266 leLeuGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValVal 282  
|||  
2951 TCCTTGAATGGGAGGATATATTTATGGACAGAAATGTCATGTTTGTGTG 3000  
283 AsnAspTrpHisAlaSerLeuValPro..Val. .... 292  
|||  
3001 AACGATTTGGCATGCCAGCCTTGTGCCAGTGTACGTTGTTTGTGGATCTGA 3050  
292 ..... 292  
3051 AAGTCCAATCCTTTATTCATTCCTGCTTTGTCAGTGTGCCCATGTCCTACA 3100  
292 ..... 292  
3101 TTCTTTTATGCTTTTTCATGCTGTTCTTATATATTGCATATATGCTTAT 3150  
292 ..... 292  
3151 GGAGTCTAAAAGTTACCGAGGGAATAACTCTTAAGGATTTTCCTCAATCA 3200  
292 ..... 292  
3201 ATTATCTTATAGCTTTTACTTAACATTTACTGTGGCAACATAATGTGTTTT 3250  
292 ..... 292  
3251 GAGATTTACAAAGTTTCAGAGATTGCACCTTCACTAGTTGCTAGCTAATCTGA 3300  
292 ..... 292  
3301 TGTTTTCCCGAGAAAATGCCTAAAGCTTTGTCTTTGATGTCATTTGATAG 3350  
292 ..... 292  
3351 AAAAAGAGTTTATGTACACTCCCAAGAGGGGACCCCAAAATTACAACACC 3400  
292 ..... 292  
3401 ACACCCCTGAGAAGTACAGGCTGCCGGAAGGATGCAAGCCCCCACTG 3450  
292 ..... 292  
3451 CCGCTGCTTATAGTCAAGCCGGGCTGAGCTTGTGTTGTCAAGTAAGC 3500  
292 ..... 292  
3501 TAGCAGTCTAGATTGCGCAAGGTCGATTCGTCGAAGATGACAGTGTTCG 3550  
292 ..... 292  
3551 GCTGCTTCCAAATCCACCAAACTATGACATGATCACTGGAGAAAGTACCT 3600  
292 ..... 292  
3601 TTTCTCGCGGTGAGGGGTGAGCTGGTGGTCTGCTGCTGCCAGTTTTC 3650  
292 ..... 292



3651 GATAATCTGAAAAATGCATGTTTTCATGATTTTAGTATCTTTGCGGACCT 3700  
292 ..... 292  
3701 GGGTACCACCTAAGCTTTTCACACAGATAATTTGCAGTTACACCTATAAAG 3750  
292 ..... 292  
3751 TAACGGTCATGATATGCATGTGTTTGGGTAGATCATGTCATGCATGCAATTT 3800  
292 ..... 292  
3801 TAGGAATTAGGACATGCCAGAACACAGTGGCTTATGGGCAATTCATT 3850  
292 ..... 292  
3851 TGTTCATTATACGAGTCATCAATATGGTTCAGCATGTTTGGACGCTACT 3900  
292 ..... 292  
3901 TGTTTGGGCAATTTTCAGATGGTGAATTTAGCTGCTTGATGTGGCTAG 3950  
292 ..... 292  
3951 CTGGCTTATTTGTACAAAGTATCGATGTTAGATGCATATTTCCCTTTTGT 4000  
292 ..... 292  
4001 CTGTGCTGTTTGGCATGTTTGTATTCCCTTTTCTGTGCGCAGTTGGCA 4050  
292 ..... 292  
4051 TGTAAATTTGTTTCAATTACATAATCAACTTTTGTGCTGACATCAGTCA 4100  
293 ..... 319  
4101 TTTTATTACGCTTCTGTGCAAAATATAGACCATTACCGTGTGTACAG 4150  
305 gAspSerArgSerThrLeuValIleHisAsnLeuAlaHisGln..... 319  
4151 AGATTCCGCGACGCCCTTGTATACATAATTTAGCACATCAGGTTTGGG 4200  
319 ..... 319  
4201 TCTATCACCTTTTCAATATCCGTACATGGCTTTGTAAAGTCGGTTTCACACGT 4250  
320 ..... 325  
4251 ATCGTCATCTGTATGTTATTTCATATGTCATTAGGGTGTGGAGCCTTGCAA 4300  
325 erThrTyrProAspLeuGlyLeuProGluTyrGlyAlaLeuGlu 341  
4301 GTACATATCTGATCTGGGATTTGCCACCTGAATGGTATGGAGCTTTAGAA 4350  
342 TrpValPheProGluTrpAlaArgHisAlaLeuAspGlyGluAl 358  
4351 TGGGTATTTCCAAATGGGCAAGAGGATGCCCTTGACAAGGGTGAGGC 4400  
358 aValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValThrV 375  
4401 AGTTAACTTTTCAAGGAGCAGTTGTGACAGCAGATCGAATGTGACCG 4450  
375 aISerGln..... 377  
4451 TCAGTCAGTGAATACTCAATACTTCTCTTTTCTTTTTCGGGATGTTTC 4500  
377 ..... 377  
4501 TTCAGTTCAATTCCTGCTTTTCCACCAATTAAGAAATGATTTAATCTT 4550  
378 ..... GlyTyrSerTrpGluValThrAlaGluGlyGlyGlnG 391  
|||||

4551 TTGTTTCTAGGGTTATTTCATGGGAGGTCACAACCTGCTCAAGGTGGACAGG 4600  
391 lyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGly.Ile... 406  
4601 GCCTCAATGAGCTCTTAAGCTCCCGAAAAAGTGTATTGAATGGTAACATAT 4650  
406 ..... 406  
4651 ATTTGAATCCACTTATCTTCTTCTGAAACATATTTTACAGAAATAGATGGA 4700  
407 ..... Val 407  
4701 TGGGTTCCAGAATAAATTTCAGTTTGGCTCTTTCGGTATGAAAGGAATTGTA 4750  
408 AsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuPr 424  
4751 AATGGAATTGCATTAATGATTGGAACCCACCACACAGACAAGTGTCTCCC 4800  
424 oHisHisTyrSerValAspAspLeuSerGlyLys..... 435  
4801 TCATCATTTATTTCTGCGATGACCTCTCTGGAAAGGTGTGTGGATAGTACC 4850  
435 ..... 435  
4851 CTATATAATAACATGATATATCTGATCTAGTACTTCTTTTCTTTTCTGCTAG 4900  
435 ..... 435  
4901 TTTGCTTCCCATGATGTTTCTCACTAACTATCTATGTTGTTTGGCATAC 4950  
436 ..... AlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProV 450  
4951 TTGTCAGGCCAAATGTAAGCTGAATTCGAGAAGGAGTGGGTTTACCTG 5000  
450 aArgGluAspValProLeu..... 456  
5001 TAAAGGAGGATGTTCTCTGTTGTAGATACAAACCCCTTAAGATATATTT 5050  
456 ..... 456  
5051 TTTAAATCCCTAAAAAACTTGCCGATCATCTCATTTAGCTTGTATTCACA 5100  
457 lLeGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleL 473  
5101 GATTGCTTTTATTGGAAGACTGGATTACCAGAAAGCATTTGATCTCATTA 5150  
473 ySMetAlaIleProGluLeuMetArgGluAspValGlnPhe..... 486  
5151 AAATGGCCATTCCAGAGCTCATGAGGGAGGAGTGCAGTTTGTAAAGTTCA 5200  
486 ..... 486  
5201 TATTCTTTTCTTGAGACTAGAGTATAAATCAACATGTAGGTGGGGT 5250  
486 ..... 486  
5251 GGTATAATACAGACATAAGTTCACGCTATTGCTTCCATGAGAATTTTAA 5300  
486 ..... 486  
5301 GCTATTTCAGTAATATGCTACTGCAAGTTTGTGAACAAAAGTTTGAAGCAAT 5350  
486 ..... 486  
5351 AAATATATGTTAGCACATGACCATGCAGTGCACCTATAGCTGGAATGTCC 5400  
486 ..... 486  
5401 TGTAAGTATGTTGATCTAAACACACTCAACACATGTTTTTTCGCATACAAAC 5450  
486 ..... 486  
5451 ACATCGGTGGCGCAACAAACATACTCTACATAAAATTTGGCTTGGTGAA 5500



486 ..... 486  
5501 CTGCAGACATGCTCTTATCTCCATTCACAACTTTCTTGTTCACATGG 5550  
486 ..... 486  
5551 CTGAAGACTAAGAAAGGGGGACCCAGGGTGATGTAGCCAACTAGATCCA 5600  
486 ..... 486  
5601 GTAAGGAAGCTAGCCGAGCCTAGGAGATTGCTTAGGTAGCTGGAACGT 5650  
486 ..... 486  
5651 AGGCTCTGACAGGGGAAGCTTCCGGGAGCTAGTCGATGCCAGTGGTGAGGA 5700  
486 ..... 486  
5701 GAGGTGTGATATCTTTTGGCTCCAGAAACCAAAATGTAGGGGACAGAAG 5750  
486 ..... 486  
5751 GCGAAGGAGGTGGAGGATACCGGCTTCAAGCTGTGGTACATGGGACGGCT 5800  
486 ..... 486  
5801 GCAACAGAAATGGCGTAGGCATCTTGATCAACAAGAGCCCTTAAGTATGG 5850  
486 ..... 486  
5851 AGTGTAGACGTCAGAGACGTGGGGACCGGATTATCCTCGTCAAGCTGG 5900  
486 ..... 486  
5901 TAGTTGGGACTTAGTTCATGTTATCAGCGTGTATGCCCGCAAGTA 5950  
486 ..... 486  
5951 GCCCACAATGAGAACGCCAAGAGGAGTTCTGGGAAGCCCTGGAAGACAT 6000  
486 ..... 486  
6001 GGTAGGAGTGTACCGATTGGCGAGAAGCTCTTCATAGAGGAGACCTCA 6050  
486 ..... 486  
6051 ATGGCCACGTGGGTACATCTAACATAGGTTTGAAGGGGCACATGGGGC 6100  
486 ..... 486  
6101 TTTGGCTATGGCATCAAGAATCAAGAAGAAGATGTCTTACGCTTTGCTCT 6150  
486 ..... 486  
6151 AGCCTACGACATGATTGTAGCTAACACCCCTCTTTAGAAAGAGAGATCAC 6200  
486 ..... 486  
6201 ATCTGTGTGACTTTTAGTAGTGCCAACTAGCCAGATCGATTTTCATGCT 6250  
486 ..... 486  
6251 CTCGAGAAGAAGATAGGTGTGCGCGCTAGACTGCAAGGTGATACCTTT 6300  
486 ..... 486  
6301 CGGATTCTGTCCAGCGGATAGCGTGGCAAGTCGCTAGATGNAAGTG 6350  
486 ..... 486  
6351 GTGGAAGCTCAAGGGGGAGGTAGCTCAGGCGTTCAAGGAGAGGGTCAITTA 6400

486 ..... 486  
6401 GGGAGGCCCTTGGAGGAAGGAGGGGATCCGACAATGTGTGATGAAG 6450  
486 ..... 486  
6451 ATGGCACTTGCAATTCGTAAAGTGGCCTCGGAGGAGTGTGGAGTGTCCAG 6500  
486 ..... 486  
6501 GGGATGGAGNAGCAAGATAAGGATACCTGGTGGTGAATGATGATGTCC 6550  
486 ..... 486  
6551 AGAAGGCAATTAAGAGAGAAAGATTGCTTTAGACGCCCTATACTTTGGAT 6600  
486 ..... 486  
6601 AGGAGTGCAGTCAACATAGAAAAGTACAAGATGGCGAAGAGGCCCCAMA 6650  
486 ..... 486  
6651 GCGAGCTGTCAAGTGAAGCAAGGGTGGGCATATGAGGATCTCTACCAAC 6700  
486 ..... 486  
6701 GGTAGGCACNAAGAAAGCGAAAGGACATCTATAGATGGCCCAAGATC 6750  
486 ..... 486  
6751 CGAGAGAGAGAAAGACAGGGGATATTGGCCCAAGTCAAAATGCATCAAGGAT 6800  
486 ..... 486  
6801 GGACAGACCAACTCTTGGTGAAGGACGAGGAGATTAAAGCATAGATGGCG 6850  
486 ..... 486  
6851 GGAGTACTTCACAAAGCTGTTCAATGGGGAGGATGAGAGTCTTACCATTG 6900  
486 ..... 486  
6901 AACTTGACGACTCCTTTGATGAGACCATCATGCGTTTATCGCGCAATC 6950  
486 ..... 486  
6951 CAGGAGTCCGAGTCAAGGAGGCTTTAAAAGGAGCAAGCGGATGGGCC 7000  
486 ..... 486  
7001 CTGATTGTATCCCATTCAGGTGTGAAAGGCCCTCGGGGACATAGCGATA 7050  
486 ..... 486  
7051 GTATGGCTAACCAAGCTATTCAACCTCATTTTCGGGCAACAAGATGCC 7100  
486 ..... 486  
7101 AGAAGATGGAGAGCAAGTATATTAGTACCATTATCAACAGAGGGGGAT 7150  
486 ..... 486  
7151 GTTCAGAGTTGTACTAATTACCATTGAATTAAGCTGATGAGCCATACAAT 7200  
486 ..... 486  
7201 GAAGCTATGGAGAGAAATCATTTGAGACCGCTTAAGAAGATGACAAGCG 7250  
486 ..... 486  
7251 TCACCAAAAAATCAGTTTGGTTTCATGCTGGAGGTCCGACCATGAAACC 7300  
486 ..... 486



[illegible]



```
595 ..... 595
9151 AAGTTTTTGCTGAGCTCTGTGTCGGTTATAGGATCGACCTTGGCTGTAGC 9200
595 ..... 595
9201 ATGCTACCTTAGTGCCCTTCTATATAGACCTTAACCTGATGGACTCACTT 9250
595 ..... 595
9251 TCTCTACATAATCATAGTAGTCGATGCGCGGAGCGCTTTTGCTTGAT 9300
595 ..... 595
9301 TCTGCTAATTTAATTTTCATGACGATAACTCATACCATGGTTTGGTCTC 9350
595 ..... 595
9351 CGATGGGGCCAGATGGCGTCTAGTGTCTGGATCTGTGAACCTAGCCA 9400
596 ..... Ala... 596
9401 ATGCGCGGTTGTTCCAAGTGAAATTTTACCTTTTGACCATTTGCGAGCA 9450
597 LeuArgThrAlaMetSerThrPheArgGluHisLysProSerTrpGluG1 613
9451 TTGCGAACCCTGATGTCGACATTCAGGAGACACAGCGCTCTGGGAGG 9500
613 yLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHisAla..Pro 629
9501 GCTCATGAAGCGGAGCATGACGAAAGACCATACGTGGGACCATGCGCGG 9550
630 SerSerThrSerArgSerSerGlyProSerTrpThrAsnProThrSe 646
9551 AGCAGTACGAGCAGATCTCGAATGGCCCTTCGTGACCAACCCACGTC 9600
646 rCysArgArgGlyLeuGlyArgSerLysCysGluSerProSerAlaLeuL 663
9601 ATGTAGACGGGACTGGGAGGTGCGAAGCGCGGTCTCTTGAGCTCTGA 9650
663 yThrSerSerSerSerPheArgGlyProGluGlyTyrProCysThrLeu 679
9651 AGACATGTTCCTCATCTTCCGCGCCCGAAGGATACCCCTGTACATTG 9700
680 ArgCysProAlaThrValGluSerGlnCysAlaCysLeuLeuTrpPheAl 696
9701 CGTTGCTCTGCTACAGTAGATGCGAATGCGCTGTGCT.TGGTCCGC 9749
696 aGlySerArgThrTyrAspGlyCysAlaAlaAlaValThrAlaSerG 713
9750 CGGTTCGAGAGTAGATGACGCTGTGCTGTCGCGCGGTGACAGCTCGG 9799
713 lyGlyArgGlnLeuGlnPheTrpGlyLleArgLysGlyCysAlaAlaGly 729
9800 GTGGATGACAGTTACAGTTTGGGGAATAGGAAGGATGCTGCTGCAGGA 9849
730 TrpLeuThrAlaLysHisSerAspGlySerLeuSerValArgValTh 746
9850 TGGTTACACCAAGACCACTCAGATGGCAGCCTCTCTGTCGCTGTAC 9899
746 rAlaGluIleArgAsnGlnLeuValThrLeu 756
9900 AGCTGAAATCAGAAACCACTGGTGACTCTT 9930

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ50642
seq_documentation_block:
ID AAZ50642 standard; cDNA; 1415 BP.
XX
AC
AC
XX
XX
DT 23-MAY-2000 (first entry)
```



67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84  
 190 CGTCCCGGGTTCCTCCG...CGCCGGCGGACCCACGGGTGAGCCG 236  
 84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100  
 237 CATCGACCGCCGCGCGTCCGACGCGCGCTGGGGACCTCGGT... 283  
 101 AspLeuLeuGluGlyIleAlaGluAspSerIleaspSerIleIleVa 117  
 284 ...CTCGAACCTGAAGGGATGCTGAAGGTCCATCGATAACACAGTAGT 330  
 117 lAlaIAserGluGlnAspSerGluIleMetAspAlaAsnGluLeuProG 134  
 331 TGTGGCAAGTGAGCAAGATTCTGAGATTGGTGGAAAGGAGCAAGCTC 380  
 134 lAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150  
 381 GAGCTAAGTAACACAAAGCATTCTCTTTGTAACCGCGGAAGCTTCCT 430  
 151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167  
 431 TATGCAAGTCTGGGGGTCTAGGAGATGTTTGTGGTTTCATTGGCCAGTGC 480  
 167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184  
 481 TCTTGTCTGCTGGGTCAACGCTGATGGTGGTGAATGCCAGATATTAA 530  
 184 sNglySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200  
 531 ATGCTACCTCCGATAAGAATTATGCAAAATGCATTTTACACAGAAACAC 580  
 201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG 217  
 581 ATTCGGATTCCATGCTTTGGCGGTGAACATGAAGTTACCTTCTTCCATGA 630  
 217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234  
 631 GTATAGAGATTTCAGTTGACTGGGTGTTTGTGATCATCCCTCATATACA 680  
 234 rGProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250  
 681 GACCTGGAAATTTATGGAGATAAGTTTGGTGTCTTTGGTGAATAACAG 730  
 251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267  
 731 TTCAGATACACACTCCTTTGCTATGCTGCATGTGAGGCTCCTTTGATCCT 780  
 267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284  
 781 TGAATTGGGAGGATATATTTATGACAGAAATTGCATGTTGTGTCATAG 830  
 284 sPTrpHisAlaSerLeuValProValLeuLeuAlaLysTyrArgPro 300  
 831 ATTGCGATGCCAGTCTAGTGCAGCTCCTCTTGTGCAAAATATAGACCA 880  
 301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317  
 881 TATGGTGTTTATAAAGACTCCCGCAGCACTTCTTGTATAACATAATTAGC 930  
 317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334  
 931 ACATCAGGGGTAGAGCCTCGAAGCACAATATCCTGACCTGGGTGGCCAC 980  
 334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350  
 981 CTGAATGGTATGAGCTCTGAGTGGGTATTCCCTGAATGGCGGAGGAGG 1030  
 351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367  
 1031 CATGCCCTTGACAAAGGTGAGGCGTAAATTTTTTGAAGGTGCAGTGT 1080

367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384  
 1081 GACAGCAGATCGAATCGTGACTGCAGTAAGGGTTATTCTGGGAGGTCA 1130  
 384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400  
 1131 CAACTGCTGAAGGTGGACAGGGCTCAATGAGCTCTTAAGCTCCAGAAAG 1180  
 401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417  
 1181 AGTGTATTAAACGGAATTGTAATGGGAATTGACATTAAATGTTGAACCC 1230  
 417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434  
 1231 TGCCACAGACAATGTATCCCTGTCAATTATCTGTTGATGACCTCTCTG 1280  
 434 lYlAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450  
 1281 GAAAGGCCAAATGTAAAGGTGCATTGCAGAGGAGCTGGGTTTACCTATA 1330  
 451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467  
 1331 AGGCCTGATGTTCTCTGATTGGCTTTATTGGAGGTTGGATTATCAGAA 1380  
 467 sGlyIleAspLeuIleLysMetalIleProGlu 478  
 1381 AGCATTTGATCTCATTCACACTTATCATACCAGAT 1414







OM of: US-09-674-824-2 to: Issued\_Patents\_NA:\* out\_format : pfs  
Date: Mar 28, 2002 5:43 PM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool/US09674824/runat\_27032002\_094251\_18439/app\_query.fasta\_1.824  
-DB=Issued\_Patents\_NA -PMF=fastap -SUFFIX=rni -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-QAPOP=4.500 -QAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRYX=blomsm62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09674824 -CGNL\_1\_94 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

## Search information block:

Query: US-09-674-824-2  
Query length: 756  
Database: Issued\_Patents\_NA.\*  
Database sequences: 351203  
Database length: 113238999  
Search time (sec): 92.300000

## score\_list:

Sequence	Strd Orig	zScore	Escore	Len	Documentation
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-196-390-1 +	3612.00	5385.06	2.7e-292	2239	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-192-909-1 +	2809.50	4180.73	3.2e-225	2383	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-12 +	2713.50	4040.48	2.1e-217	1752	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-572-951-1 +	2651.00	3940.26	8.0e-212	2990	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-10 +	2001.00	2968.27	1.1e-157	2360	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-836-567-9 +	1731.50	2567.65	2.3e-135	1758	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-10 +	1246.00	1837.36	1.1e-94	2097	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-836-567-7 +	1215.00	1787.39	6.6e-92	2793	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-836-567-5 +	1184.00	1745.11	1.4e-89	1926	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-196-390-5 +	1182.50	1738.51	3.5e-89	2825	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-8 +	1076.00	1582.93	1.6e-80	2007	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-572-951-2 +	1071.00	1574.97	4.5e-80	2085	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-572-951-3 +	943.50	1382.14	2.5e-69	2380	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-6 +	900.00	1316.10	1.2e-65	2542	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-07-735-065-1 +	678.50	990.59	1.6e-47	1464	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-469-202-11 +	678.50	990.59	1.6e-47	1464	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-735-491-1 +	634.50	923.52	8.6e-44	1601	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-470-720-2 +	612.50	884.88	1.2e-41	2549	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-470-720-5 +	604.50	864.80	1.6e-40	4964	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-4 +	603.00	862.96	2.0e-40	4800	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-115-704-1 +	558.50	803.41	4.2e-37	2652	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-08-836-567-1 +	556.00	801.37	5.5e-37	2303	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-836-567-11 +	556.00	794.18	1.4e-36	4168	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-968-542C-1 +	526.00	744.71	7.9e-34	6027	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-194-905-7 +	157.50	190.48	0.0058	6854	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-370-700-1 +	153.00	153.91	0.6355	80161	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-08-804-227C-7 +	152.00	159.58	0.3071	44377	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-804-198-1 +	148.50	159.58	0.3071	44377	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-07-755-573C-7 +	146.50	148.91	0.0042	1810	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-09-428-517-1 +	146.00	148.91	1.21	50937	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-07-642-734C-3 +	142.50	154.86	0.5628	20235	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-938-009A-3 +	142.50	154.86	0.5628	20235	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-938-291A-3 +	140.50	168.76	0.0946	5020	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-804-227C-7 +	140.50	142.33	2.80	44377	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-804-198-1 -	140.50	142.33	2.80	44377	
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-443-501A-2 +	140.00	135.72	6.55	71989	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-097-997A-12 +	137.00	167.68	0.1087	3561	

/cgn2\_6/ptodata/1/ina/6A\_COMB.seq:US-08-665-574C-12 + 137.00 167.68 0.1087 3561  
/cgn2\_6/ptodata/1/ina/6B\_COMB.seq:US-08-946-994-12 + 137.00 167.68 0.1087 3561  
/cgn2\_6/ptodata/1/ina/5A\_COMB.seq:US-08-614-770A-1 - 137.00 135.82 6.47 49272  
/cgn2\_6/ptodata/1/ina/6A\_COMB.seq:US-08-658-136-2 + 137.00 134.81 7.36 53526  
/cgn2\_6/ptodata/1/ina/6A\_COMB.seq:US-08-658-136-1 + 137.00 134.80 7.37 53577  
  
seq\_name: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:US-09-196-390-1  
  
seq\_documentation\_block:  
: Sequence 1, Application US/09196390  
: Patent No. 6307125  
: GENERAL INFORMATION:  
: APPLICANT: Block, Martina  
: APPLICANT: Lorz, Horst  
: APPLICANT: Lutticke, Stephanie  
: APPLICANT: Walter, Lennart  
: APPLICANT: Froberg, Claus  
: APPLICANT: Kossmann, Jens  
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES  
: TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH  
: NUMBER OF SEQUENCES: 9  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave  
: STREET: 1251 Avenue of the Americas  
: CITY: New York  
: STATE: New York  
: COUNTRY: United States of America  
: ZIP: 10020  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: OPERATING SYSTEM: IBM PC compatible  
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/196,390  
: FILING DATE:  
: CLASSIFICATION:  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: DE 196 21 588.9  
: FILING DATE: 29-MAY-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: DE 196 36 917.7  
: FILING DATE: 11-SEP-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: PCT/EP97/02793  
: FILING DATE: 28-MAY-1997  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Haley, Jr., James F.  
: REGISTRATION NUMBER: 27,794  
: REFERENCE/DOCKET NUMBER: AGREVO-9  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (212) 596-9000  
: TELEFAX: (212) 596-9090  
: INFORMATION FOR SEQ ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 2239 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: cDNA to mRNA  
: HYPOTHETICAL: NO  
: ANTI-SENSE: NO  
: ORIGINAL SOURCE:  
: ORGANISM: Triticum aestivum L.  
: STRAIN: cv. Florida  
: HAPLOTYPE: ca. 21 d Caryopses  
: IMMEDIATE SOURCE:  
: LIBRARY: cDNA library in pBluescript sk (-)  
: CLONE: Tasss  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: 3..2017



Fri Mar 29 09:26:44 2002

US-09-196-390-1

alignment\_scores:  
Quality: 3612.00 Length: 671  
Ratio: 5.383 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-674-824-2 x US-09-196-390-1 ..

Align seg 1/1 to: US-09-196-390-1 from: 1 to: 2239

86 ThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaProAspLe 102  
3 ACGCAGCGCCCTGCGGACGCCGGGTGGGGAACCTCGCGCCGACCT 52  
102 uLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleValAlaA 119  
53 CCGTCTCGAAGGGATTGCTGAGGATTCATCGACAGCATAAATTGGCTG 102  
119 laSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProGlnAla 135  
103 CAAGTGACAGGATTCAGATCATGGATGCGAATGAGCAACCTCAAGCT 152  
136 LysValThrArgSerIleValPheValThrGlyGluAlaAlaProTyrAl 152  
153 AAAGTTACAGTAGCATCGTCTTTGTGACTGGTGAAGCTGCTCTTATGC 202  
152 alySerGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuA 169  
203 AAAGTCAGGGGGTGGGAGATGTTTGGTTCGTTACCAATGCTCTTG 252  
169 laAlaArgGlyHisArgValMetValMetProArgTyrLeuAsnGly 185  
253 CTCTCGTGTGCCAGGATGATGGTGTGAATGCCAAGATACTTAATGGG 302  
186 SerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLy 202  
303 TCCTCTGATAAAACATATCAAGGGATTTATACACTGCGAAGCACATTAA 352  
202 sileProCysPheGlyGlySerHisGluValThrPhePheHisGluTyrA 219  
353 GATTCATGCTTTGGGGGATCACATCAAGTGACCTTTTTCATGAGTATA 402  
219 rgAspAsnValAspTrpValPheValAspHisProSerTyrHisArgPro 235  
403 GAGACAACGTCGATTGGGTGTTTGTGATCATCCGTCATATCACAGACCA 452  
236 GlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheAr 252  
453 GGAAGTTTATATCGAGATAATTTTGGTGTCTTTTGGTGATAATCAGTTCA 502  
252 gTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluL 269  
503 ATACACACTCCTTTGCTATGCTGATCGAGGCCGCCCACTAACTCTTGAAT 552  
269 euGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrp 285  
553 TGGGAGGATATATTTATGGACAGANTTGCATGTTTGTGTGAACGATTGG 602  
286 HisAlaSerLeuValProValLeuLeuAlaLysTyrArgProTyrG1 302  
603 CATGCCAGCCTTGTGCCAGTCTCTTGTGCTGCAAAATATAGACNCATACGG 652  
302 yValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisG 319  
653 TGTTTACAGATTCGCCGACGACCCCTTGTATACATAAATTACACATC 702  
319 lnglyValGluProAlaSerThrTyrProAspLeuGlyLeuProProGlu 335  
703 AGGGTGTGAGCCCTGCAAGTATACATATCTGTGCTGGGATTTGCCCTCTGAA 752

336 TrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAl 352  
753 TGGTATGGAGCTTTAGAAATGGTATTTCCAGAAATGGGCAAGGAGCATGC 802  
352 aLeuAspLysGlyGlyAlaValAsnPheLeuLysGlyAlaValValThrA 369  
803 CCTTGACAAGGGTGGAGGAGTTAACTTTTGAAGGAGCAGTGTGGACAG 852  
369 laAspArgIleValThrValSerGlnGlyTyrSerTrpGluValThrThr 385  
853 CAGATCGGATTGTCACCGTCAGTCAGGTTTATTCATGGAGGTCAAACT 902  
386 AlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerVa 402  
903 GCTGAAGGTGGACAGGGCCTCAATGAGCTCTTAAGCTCCCGAAAAAGTGT 952  
402 lLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrT 419  
953 ATTGAATGGAATTTGTAATGGAAATTGACATTAATGATTGGAAACCCACCA 1002  
419 hrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerGlyLys 435  
1003 CAGACAAGTGTCTCCCTCATCATTAATTCCTCGATGACCTCTCTGGAAG 1052  
436 AlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgG1 452  
1053 GCCAAATGTAAGCTGAATTCAGAAAGGAGTTGGGTTTACCTGTAAAGGA 1102  
452 uAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyI 469  
1103 GGATGTTCTCTGATGGCTTTATTTGAAGACTGGATTACCAAGAAAGGCA 1152  
469 leAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGln 485  
1153 TTGATCTCATTAATAATGGCCATTCACAGCTCATGAGGAGGACGTGCAG 1202  
486 pheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgSe 502  
1203 TTTGTCATGCTTGGATCTGGGGATCCAATTTTCAAGGCTGGATGAGATC 1252  
502 rThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerV 519  
1253 TACCGAGTCGAGTTACAAGGATAAATTCCTGGATGGGTTGGATTTAGTG 1302  
519 alProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetPro 535  
1303 TTCAGTTTCCACAGAAATAACTGCAGGTTCCGATATATTTGTTAATGCCA 1352  
536 SerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrG1 552  
1353 TCGAGATTGAACTTGGGGTCTTAATCAGCTATATGCTATGCAATATGG 1402  
552 yThrValProValValHisGlyThrGlyGlyLeuArgAspThrValGluT 569  
1403 TACAGTTCTGTAGTTTCATGGAACCTGGGGGCTCCGAGACACACATCGAGA 1452  
569 hrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPhe 585  
1453 CCTTCAACCCCTTTGGTGCAGAAAGAGAGAGGAGGTACAGGTTGGCGGCTC 1502  
586 SerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSe 602  
1503 TCACCGCTAACCGTGCAGCAAGATGTTTGGGCATTTCCGAACCGCATGTC 1552  
602 rThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyM 619  
1553 GACATTCAGGAGACACACCCCTCTCTGGAGGGGCTCATGAAGCGAGGCA 1602  
619 etThrLysAspHisThrTrpAspHisAlaProSerSerThrSerArgSer 635  
1603 TGACGAAAGACCATACCTGGGACCATGCCCCGAGCAGTACGAGCAGATCT 1652  
636 SerSerGlyProSerTrpThrAsnProThrSerCysArgArgGlyLeuG1 652



```

|||||
1653 TCGAGTGGCCCTCGTGGACCAACCTACGTACGTACGTACGGGACCTGGG 1702
|||||
652 YArgSerIscysGluSerProSerAlaLeuLysThrSerSerSerP 669
|||||
1703 GAGGTCCAAGTGGCAGCTCTCTCAGCTCTGAAGACATCTCTTCATCT 1752
|||||
669 heArgGlyProGluGlyTyrProCysThrLeuArgCysProAlaThrVal 685
|||||
1753 TCCGCGCCCGAAGATACCCCTGTACATGCGTTGCTCTGTACAGTA 1802
|||||
686 GluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrTyrAs 702
|||||
1803 GAGTCGCAATGCGCTCTGCTTGTGTTGCTGCGGTTTCGAGAACATATGA 1852
|||||
702 pGlyCysAlaAlaAlaValThrAlaSerGlyArgGlnLeuGlnP 719
|||||
1853 CGGCTGTGCTGCTGCGCGGTGACAGTTCGGGTGGACAGTTACAGT 1902
|||||
719 heTrpGlyTleArgLysGlyCysAlaAlaGlyTrpLeuThrAlaLysHis 735
|||||
1903 TTTGGGAATAAGGAGGATGTCTGCAGGATGTTAACAGCAACAC 1952
|||||
736 HisSerAspGlySerLeuSerValArgValThrAlaGluLleArgAsnG 752
|||||
1953 CACTCAGATGCGAGCTCTCTGTCGCTGTACAGCTGAATCAGAAACCA 2002
|||||
752 nLeuValThrLeu 756
|||||
2003 ACTGGTGACTCTT 2015

```

seq\_name: /cqn2.6/ptodata/1/1na/6B\_COMB.seq:us-09-192-909-1

```

seq_documentation_block:
; Sequence 1, Application US/09192909
; Patent No. 6307124
; GENERAL INFORMATION:
; APPLICANT: Jens Kossmann
; APPLICANT: Claus Frohberg
; TITLE OF INVENTION: Nucleic acid molecules encoding soluble
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/192,909
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02527
; FILING DATE: 16-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 19 918.2
; FILING DATE: 17-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: GFB-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-96-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 2383 base pairs
; TYPE: nucleotide
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Zea mays
; TISSUE TYPE: endosperm
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1950
; OTHER INFORMATION: /function= "starch synthesis"
; OTHER INFORMATION: /product= "soluble starch synthase"
US-09-192-909-1

alignment_scores:
    Quality: 2809.50      Length: 749
    Ratio: 4.411         Gaps: 10
    Percent Similarity: 85.047      Percent Identity: 73.565

alignment_block:
US-09-674-824-2 x US-09-192-909-1 ...

Align seg 1/1 to: US-09-192-909-1 from: 1 to: 2383

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20
   ::::|||||:|||||
41 TCGGCGGTGGCGCGCGCTGCTCTCCTC..... 67

20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34
   ||| ||||| :||| |||||
68 .CTCGCGCGCGCGCGCTGCGCGCGCGCTGCGCGCGCGCGCGCGCG 116

35 ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
   |||||:|||||: ||| ||| ||||| |||||
117 GCGCGCTCCAGCGCGTGTGCGCGCGCGTGTGCGCGAGTGACGAG 166

51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
   ||||| ||||| |||||
167 GAGGCGCGCGCGCGCGCGATGCCACCGCGCTGCTGCGCGCGCGCT 216

67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84
   ||||| ||||| ||||| |||||
217 CGTGCGCGCTTCTCTCGCG...CCGCGCGCGCGCGCGCGGTGAGCGG 263

84 laProThrGlnProProLeuProAspAlaGlyValGlyLeuAlaPro 100
   || ||| |||||:|||||:||||| |||||
264 CATTGACGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGGTCTCTCGGTGTC 313

101 AspLeuLeuLeuGlnGlyTleAlaGluAspSerIleAspSerIleLeu 117
   :||| ||||| ||||| |||||
314 GAA.....CCTGAAGGGATTGCTGAAGGTTCCATCGATAACACAGTAGT 357

117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnPro 134
   l:||||| ||||| ||||| |||||
358 TGTGCAAGTGAGCAAGATTCTGAGATTGTGTTGGAAGGAGCAAGCTC 407

134 lAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
   :||| |||||:|||||:||||| |||||
408 GAGCTAAAGTAACACAAAACATTGCTTTGTAACCTGCGAAGCTTCTCT 457

151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
   ||||| ||||| ||||| |||||
458 TATGCAAGTCTGGGGGTCTAGGAGATTGTTGTTGTTTTCATTGCCAGTTGC 507

167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeu 184
   ||||| ||||| ||||| |||||
508 TCTTGCTCTGTCACCGTGTGATGTTGTAATGCCAGATATTATAA 557

184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200

```



```
|||||.....|||||.....|||||.....|||||.....|||||
558 ATGTCACCTCCGATAACAATATGCAATGCAATTTACACAGAAAACAC 607
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG1 217
|||||.....|||||.....|||||.....|||||.....|||||
608 ATTCGGATTCCATGCTTTGGCGGTGAACATGAAGTTACCTTCTTCCCATGA 657
217 uTyraGAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
|||||.....|||||.....|||||.....|||||.....|||||
658 GTATAGAGATTTCAGTTGACCTGGGTGTTGTTGATCATCCCTCATATCACA 707
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
|||||.....|||||.....|||||.....|||||.....|||||
708 GACCTGGAATTTATATGGAGATAAGTTTGGTGCTTTGGTGATAATCAG 757
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267
|||||.....|||||.....|||||.....|||||.....|||||
758 TTCAGATACACACTCCTTTGCTATGCTGCATGTGAGGCTCCTTTGGTCC 807
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284
|||||.....|||||.....|||||.....|||||.....|||||
808 TGAATGGGAGGATATATTTATGGACAGAAATGCTGTTGTTGTTGCTCAATG 857
284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
|||||.....|||||.....|||||.....|||||.....|||||
858 ATTGGCATGCCAGTCTAGTGCACGCTCTTCTTGTGCAAAATATAGACCA 907
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317
|||||.....|||||.....|||||.....|||||.....|||||
908 TATGGTGTATAAAGACTCCCGGAGCATCTCTTGTATACATAATTTAGC 957
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334
|||||.....|||||.....|||||.....|||||.....|||||
958 ACATCAGGGTCTAGAGCCTGCAACACATATCCTGACCTGGGTGGCCAC 1007
334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350
|||||.....|||||.....|||||.....|||||.....|||||
1008 CTGAATGATATGAGCTCTGGAGTGGGTATTTCCCTGAATGGCGGAGGAG 1057
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367
|||||.....|||||.....|||||.....|||||.....|||||
1058 CATGCCCTTGACAGGGTGAGCGACTTAATTTTGAAGGTGCAGTTGT 1107
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384
|||||.....|||||.....|||||.....|||||.....|||||
1108 GACAGCAGATCGAATCGTGACTGTGCTGCTGAGTAAAGGTATTCATGGGAGTCA 1157
384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
|||||.....|||||.....|||||.....|||||.....|||||
1158 CAACCTGCTGAAGGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCAGAAAG 1207
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417
|||||.....|||||.....|||||.....|||||.....|||||
1208 AGTGATTAAACGGGAATGTAATGGAATGACATTAAATGATTGGAACCC 1257
417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434
|||||.....|||||.....|||||.....|||||.....|||||
1258 TGCCACAGACAAATGATATCCCTGTGCTTATTTCTGTTGATGACCTCTCTG 1307
434 lYsAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
|||||.....|||||.....|||||.....|||||.....|||||
1308 GAAAGCCCAATGTAAAGGTGCATTTGCAGAGGAGCTGGGTTCCTATA 1357
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
|||||.....|||||.....|||||.....|||||.....|||||
1358 AGGCCGTGATGTTCTCTGATGCTGCTTATTTGGAAGATTGGATTATCAGAA 1407
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484
|||||.....|||||.....|||||.....|||||.....|||||
1408 AGGCATTTGATCTCATTCACATTTATCAACCATCTCATGCCGGGAGATG 1457
484 aIcInPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
|||||.....|||||.....|||||.....|||||.....|||||
```

seq\_name: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq;US-08-941-445A-12

seq\_documentation\_block:  
; Sequence 12, Application US/08941445A  
; Patent No. 6107060  
; GENERAL INFORMATION:  
; APPLICANT: Keeling, Peter  
; APPLICANT: Guan, Hanping  
; TITLE OF INVENTION: Starch Encapsulation  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:



```

: ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
: STREET: 5370 Manhattan Circle
: CITY: Boulder
: STATE: CO
: COUNTRY: US
: ZIP: 80303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/941,445A
: FILING DATE: 30-SEP-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/026,855
: FILING DATE: 30-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Winner, Ellen P
: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 89-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1752 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: not relevant
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1752
: US-08-941-445A-12

```

```

alignment_scores
  Quality: 2713.50      Length: 585
  Ratio: 4.916          Gaps: 3
  Percent Similarity: 94.359      Percent Identity: 86.325

alignment_block
US-09-674-824-2 x US-08-941-445A-12

Align seg 1/1 to: US-08-941-445A-12 from: 1 to: 1752

45 ValAlaGluLeuSerArgGluGlyProAlaAlaArgProAlaGlnGlnG1 61
|||||
4 GTCGGGAGCTGAGCAGGAGGGGCCCGCCCGCCCGCTGCCACCCGC 53

61 nGlnLeuAlaProProLeuValProGlyPheLeuAlaProProProA 78
|||||
54 GCTGCTGGCCCGCCCGCTCGTGGCCGCTTCCTCGCG...CCGCCGCGC 100

78 laProAlaGlnSerProAlaProThrGlnProProLeuProAspAlaGly 94
|||||
101 AGCCACGGGTGAGCGGCATCCACCGCCCGCCCGCTGCCACCGCGC 150

95 ValGlyGluLeuAlaProAspLeuLeuLeuGluGlyLeuAlaGluAspSe 111
|||||
151 CTGGGGGACCTCGGT.....CTCGAACCTGAAGGGATTCTCAAGGTTTC 194

111 rIleAspSerIleIleValAlaAlaSerGluGlnAspSerGluIleMeta 128
|||||
195 CATCGATACACAGTAGTGTGGCAAGTGGCAAGATTCTGAGATTGTGG 244

128 spAlaAsnGluGlnProGlnAlaLysValThrArgSerIleValPheVal 144
|||||

```

```

245 TTGGAAGGAGCAAGCTCGAGCTAAAGTAACACAAAGCAATTGTCTTTGTA 294
145 ThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuGlyAspValCy 161
|||||
295 ACCGGCGAAGCTTCTCCTTATGCAAGTCTGGGGGTCTAGGAGATGTTG 344
161 sGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValV 178
|||||
345 TGGTTTCATTGCCAGTCTGCTTGTCTGCTGCTGGTGGTCAACGTTG 394
178 alMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAla 194
|||||
395 TAATGCCAGATATTAAATGGTACCTCCGATAAGAATTATGCAAAATGCA 444
195 LeuTyrThrAlaLysHisIleLysIleProCysPheGlyGlySerHisG1 211
|||||
445 TTTTACACAGAAAAACACATTCGGATTCCTTGGCGGTGAACATGA 494
211 uValThrPhePheHisGluTyrArgAspAsnValAspTrpValPheValA 228
|||||
495 AGTTACCTTCTCCATGAGTATAGATTCAGTTGACTGGGTGTTGTTG 544
228 sPHisProSerTyrHisArgProGlySerLeuTyrGlyAspAsnPheGly 244
|||||
545 ATCATCCTCATATCACAGACCTGGAATTTATATGAGATAGATTGGT 594
245 AlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCystTyrAlaAlaCy 261
|||||
595 GCTTTTGGTGATAATCAGTTTCAGATACACACTCCTTTGCTATGCTGCATG 644
261 sGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGlnAsnC 278
|||||
645 TGAGGCTCCTTTGATCCTTGAATGGGAGGATATATTTATGGACAGAA 694
278 ysMetPheValValAsnAspTrpHisAlaSerLeuValProValLeuLeu 294
|||||
695 GCATGTTTGTGTCATGATTGGCATGCCAGTCTAGTCCAGTCCCTCTT 744
295 AlaAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLe 311
|||||
745 GCTGCAAAATATAGACCATATGTTGTTTATAAAGACTCCCGCAGCATTC 794
311 uValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrP 328
|||||
795 TGTAAATACATAATTTAGCACATCAGGGGTGTAGAGCCTGCAAGCACATATC 844
328 roAspLeuGlyLeuProProGluTyrTyrGlyAlaLeuGluTrpValPhe 344
|||||
845 CTGACCTTGGGTGGCCACCTGAATGCTATGGAGCTCTGGAGTGGGTATTTC 894
345 ProGluTrpAlaArgHisAlaLeuAspLysGlyGluAlaValAsnPh 361
|||||
895 CCTGAATGGCGGAGGAGCATGCCCTTGACAAGGGTGAGGCAGTTAATTT 944
361 eLeuLysGlyAlaValThrAlaAspArgIleValThrValSerGlnG 378
|||||
945 TTTGAAGGTGAGTGTGTACAGCAGATGAATCGTACTGTGCTGCTAAGG 994
378 lyTyrSerTrpGluValThrThrAlaGluGlyGlnGlyLeuAsnGlu 394
|||||
995 GTTATTCTGGGAGGTCAACAACCTGCTGAAGGTGGACAGGGCCTCAATGAG 1044
395 LeuLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAs 411
|||||
1045 CTCCTTAAGCTCCAGAAAGAGTGTATTAACGGAATTTGTAATGGAATTGA 1094
411 pileAsnAspTrpAsnProThrThrAspLysCysLeuProHisHisTyrS 428
|||||
1095 CATTAATGATTGGAACCTGCCACAGACAAATGTATCCCTGCTCATATT 1144
428 ervAlaAspAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGlnLys 444
|||||
1145 CTGTTGATGACCTCTCTGAAAGGCCAAATGTAAAGGTGCAATTGCGAAG 1194

```







```
138 hrArgSerIleValPheValThrGlyGluAlaPro.TyrAlaLysSe 154
|||||..... |||||..... |||||..... |||||..... |||||
992 CACAAAGCATTCTTTGTAACCGCGAAGCTTCTCTTAATCGAAGTC 1041
154 rGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuAlaA 171
|||||..... |||||..... |||||..... |||||..... |||||
1042 TGGGGGCTCTAGAGAGATGTTTGTGGTTCATTGCCAGTCTGCTCTGTGCTC 1091
171 rGlyHisArgValMetValValMetProArgTyrLeuAsnGlySer 187
|||||..... |||||..... |||||..... |||||..... |||||
1092 GCGGTCACCGTGTGATGTTGTAATGCCAGACATTTAAATGGTACCTCC 1141
188 AspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIlePr 204
|||||..... |||||..... |||||..... |||||..... |||||
1142 GATAAGAAATTATGCAATGCAATTTTACTCAGAAAACACATTCGGATTCC 1191
204 ocysPheGlyGlySerHisGluValThrPhePheHisGluTyrArgAspA 221
|||||..... |||||..... |||||..... |||||..... |||||
1192 ATTCCTTTGGCGGTGAACATGAAGTTACCTTCTTCATGAGTATAGAGATT 1241
221 snValAspTrpValPheValAspHisProSerTyrHisArgProGlySer 237
|||||..... |||||..... |||||..... |||||..... |||||
1242 CAGTTAGTGGGTGTTGTTGATCATCCCTCATATCAGACACCTGGAAAT 1291
238 LeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrTh 254
|||||..... |||||..... |||||..... |||||..... |||||
1292 TTATATGGAGATAAGTTTGGTGTCTTTGGTGATAATCAGTTTCAGATACAC 1341
254 rLeuLeuCysTyrAlaAlaCysGluAlaProIleLeuLeuGlyG 271
|||||..... |||||..... |||||..... |||||..... |||||
1342 ACTCCTTTGCTATGCTGATGTGAGGCTCCTTTGGTCTTGAATTTGGGAG 1391
271 lyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrpHisAla 287
|||||..... |||||..... |||||..... |||||..... |||||
1392 GATATATTATGGACAGAAATTCATGTTGGTTGTCATGATGGCATGCC 1441
288 SerLeuValProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTy 304
|||||..... |||||..... |||||..... |||||..... |||||
1442 AGTCTAGAGCCAGCTCTTCTTGTCTGCAAAATATAGACCATATGCTGTTTA 1491
304 rArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyV 321
|||||..... |||||..... |||||..... |||||..... |||||
1492 TAAAGACTCCCGCAGCATTCTTGTAAATACATAATTTAGCACATCAGGGTG 1541
321 alGluProAlaSerThrTyrProAspLeuGlyLeuProProGluTrpTyr 337
|||||..... |||||..... |||||..... |||||..... |||||
1542 TAGAGCCTGCAAGCACATATCCTGACCTTGGGTTGCCACCTGAATGCTAT 1591
338 GlyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAlaLeuAs 354
|||||..... |||||..... |||||..... |||||..... |||||
1592 GGACCTCTGGAGTGGGTATTCCTTGAATGGCGGAGGAGGATGCCCTTGA 1641
354 pLysGlyGluAlaValAsnPheLeuLeuLysGlyAlaValValThrAlaAspA 371
|||||..... |||||..... |||||..... |||||..... |||||
1642 CAAGGCTGAGGACGTTAAATTTTGAAGGTGCAGTTGTCACAGCAGATC 1691
371 rGlyValThrValSerGlnGlyTyrSerTrpGluValThrThrAlaGlu 387
|||||..... |||||..... |||||..... |||||..... |||||
1692 GAATCGTGACTGTGACGTAAGGGTATTTCATGGGAGGTACAACTGCTGAA 1741
388 GlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAs 404
|||||..... |||||..... |||||..... |||||..... |||||
1742 GGTGGACAGGCCCTCAATGAGCTTAAAGCTCCAGAAAGAGTGTATATAA 1791
404 nGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspL 421
|||||..... |||||..... |||||..... |||||..... |||||
1792 CGGAATTTGTAATGGAATTGACATTAATGATTGGAACCTTGCACACAGACA 1841
421 yscysLeuProHisHisTyrSerValAspAspLeuSerGlyLysAlaLys 437
|||||..... |||||..... |||||..... |||||..... |||||
1842 AATGTATCCCTCTGCTATTCTTGTGTGATGACCTCTCTTGAAGAGGTAAA 1891
438 cysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAspVa 454
```

```
|||||..... |||||..... |||||..... |||||..... |||||
1892 TGTAAAGGTGCATTGCAGAAGGAGCTGGGTTTACCTATAAGGCCTGTGATG 1941
454 lProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspL 471
|||||..... |||||..... |||||..... |||||..... |||||
1942 TCCCTGATTGGCTTATTGGAAGATTGGAATTATCAGAAAGGCATTGATC 1991
471 euIleLysMetAlaIleProGluLeuMetArgGluAspValGlnPheVal 487
|||||..... |||||..... |||||..... |||||..... |||||
1992 TCATTCAACTTATCATACCATCTCATCGGGAAGATGTTCAATTTGTC 2041
488 MetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgSerThrGl 504
|||||..... |||||..... |||||..... |||||..... |||||
2042 ATGCTTGGATCTGGTGACCCAGAGCTTGAAGATTGGATGAGATCTACAGA 2091
504 uSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerValProV 521
|||||..... |||||..... |||||..... |||||..... |||||
2092 GTCGATCTTCAAGCATAAATTTCTGGATGGGTTGGATTAGTGTTCAG 2141
521 alSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArg 537
|||||..... |||||..... |||||..... |||||..... |||||
2142 TTTCCACCCGGAATAACTCGGCTGGCGATATATTGTTAATGCCATCCAGA 2191
538 PheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrVa 554
|||||..... |||||..... |||||..... |||||..... |||||
2192 TTCGAACCTTGTGCTCTCAATTCAGCTATATGCTATGCAATGTCACAGT 2241
554 lProValValHisGlyThrGlyGlyLeuArgAspThrValGluThrPheA 571
|||||..... |||||..... |||||..... |||||..... |||||
2242 TCCGTGTTGCCATGCAACTGGGGCCCTTAGAGATACCGTGGAGAACTCA 2291
571 snPropheGlyAlaLysGlyGluGlyThrGlyTrpAlaPheSerPro 587
|||||..... |||||..... |||||..... |||||..... |||||
2292 ACCCTTTCGGTGAGAAATGGAGCAGGTACAGGTGGGCATTCGCACCC 2341
588 LeuThrValAspLysMet.LeuTrpAlaLeuArgThrAlaMetSerThrP 604
|||||..... |||||..... |||||..... |||||..... |||||
2342 CTAACACACAAACATGTTTGTGGACATTCGCAACTGCAATATCTACAT 2391
604 heArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMetThr 620
|||||..... |||||..... |||||..... |||||..... |||||
2392 ACAGGAACACAAAGTAATAATGGAGGGCTAATGAAGCCAGGCATGTCA 2441
621 LysAspHisThrTrpAspHisAlaPro.SerSerThrSerArgSerSerS 637
|||||..... |||||..... |||||..... |||||..... |||||
2442 AAGAGTTCACGTGGGACCATCGCGTGAACAATACGAACAAATCTTCCA 2491
637 erGlyProSerTrpTrpAsnPro 644
|||||..... |||||..... |||||
2492 GTGGGCTTTCATCGGATCGACCC 2514
```

seq\_name: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:US-08-941-445A-20

```
seq_documentation_block:
; Sequence 20, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```



```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/941,445A
: FILING DATE: 30-SEP-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/026,855
: FILING DATE: 30-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Wimer, Ellen P
: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 89-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1620 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: not relevant
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHEetical: NO
: NAME/KEY: CDS
: LOCATION: 1..1620
: US-08-941-445A-20

alignment_scores:
  Quality: 2546.50      Length: 538
  Ratio: 4.935          Gaps: 3
  Percent Similarity: 95.911 Percent Identity: 87.918

alignment_block:
US-09-674-824-2 x US-08-941-445A-20 ..

Align seg 1/1 to: US-08-941-445A-20 from: 1 to: 1620

95 ValGlyGluLeuAla...ProAspLeuLeuLeu.....GluGlyIleAl 108
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4 GTCGGGAGCTGAGCAGGAGGAGCCTCGCTCGAACCTGAAGGATGC 53

108 aGluAspSerIleAspSerIleIleValAlaAlaSerGluGlnAspSerG 125
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
54 TGAAGTTCCATCGATAACACAGTAGTTGTGGCAAGTGCAGCAAGATTCTG 103

125 luIleMetAspAlaAsnGluGlnProGlnAlaLysValThrArgSerIle 141
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
104 AGATTGTGGTTGGAAAGGAGCAAGCTCGAGCTAAAGTAACACAAAGCAAT 153

142 ValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuGl 158
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
154 GTCTTGTAAACCGGGAAGCTTCTCTTATGCAAGTCTGGGGGTCTAGG 203

158 yAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgV 175
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
204 AGATGTTGTGCTTCATTCGCCAGTTGCTCTTCTGCTGCTGCCCTGACC 253

175 alMetValValMetProArgTyrLeuAsnGlySerSerAspLysAsnTyr 191
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
254 TGATGGTTGTAATGCCAGATATTAAATGGTACCTCCGATCAAGATATAT 303

192 AlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheGlyGl 208
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
304 GCAATGCAATTTTACACAAAAACACATTTCGGATTCCATGCTTTGGCGG 353

208 ySerHisGluValThrPhePheHisGluTyrArgAspAsnValAspTpv 225
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
354 TGAACATGAAGTACCTTCTCCATGAGTATAGACATTCCAGTTGACTGGG 403

225 alPheValAspHisProSerTyrHisArgProGlySerLeuTyrGlyAsp 241
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
404 TGTGTTGTCATCATCCCTCATATACAGACCTGGAAATTTATATGAGAT 453

```



```

542 GlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValValHi 558
|||||
1354 GGTCTCAATCAGCTATATGCTATGAGTATGCCACAGTTCCTGTGTGCTCA 1403
|||||
558 sGlyThrGlyLeuArgAspThrValGluThrPheAsnProPheGlyA 575
|||||
1404 TGCAACTGGGGCCCTAGAGATACCGTGGAGAACTTCAACCTTTCGGGTG 1453
|||||
575 laLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrValAsp 591
|||||
1454 AGAATGGAGACAGCGGTACAGGGTGGGCATTCGCCACCCCTAACACAGAA 1503
|||||
592 LysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisL 608
|||||
1504 AACATGTTTGTGGACATTCGCAACTGCCAATATCTACATACAGGAAACACA 1553
|||||
608 ysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisThr 624
|||||
1554 AGTCCTCTCTGGGAAGGGCTAATGAAGCGAGGCATGTCAAAAGACTTCACG 1603
|||||
625 TrpAspHisAla 628
|||||
1604 TGGGACCACGCC 1615

```

seq\_name: /cgn2\_5/ptodata/1/ina/6A\_COMB.seq:us-08-836-567-9

```

seq_documentation_block:
; Sequence 9, Application us/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Koosmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2360 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

```

```

; ORIGINAL SOURCE:
; ORGANISM: Solanum tuberosum
; STRAIN: cv. Dshire
; TISSUE TYPE: leaf tissue
; IMMEDIATE SOURCE:
; LIBRARY: cDNA-library in Lambda ZAPII
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 68..1990
; US-08-836-567-9

```

```

alignment_scores:
Quality: 2001.00 Length: 545
Ratio: 4.222 Gaps: 5
Percent Similarity: 86.972 Percent Identity: 66.789
alignment_block:
US-09-674-824-2 x US-08-836-567-9
Align seg 1/1 to: US-08-836-567-9 from: 1 to: 2360
102 LeuLeuLeuGluGlyIleAlaGluAspSer.....IleAspSerIleI 116
|||||
335 TTGATTCACATCTCTGTGCAGGAGATGCAACAATGGTAGAATCTCATGA 384
|||||
116 eValAlaAsnSerGluGlnAsp.....SerGluIleMetAspAlaAsnG 131
|||||
385 TATTGTAGCCAATCATAGAGATGACTTGAGTGAGGATAGGAGATGG 434
|||||
131 luGlnProGlnAlaLysValThrArgSerIleValPheValThrGlyGlu 147
|||||
435 AGGAACCCCAATCAATAATCAATCAATCAATCAATCAATCAATCAATCA 484
|||||
148 AlaAlaProTyrAlaLysSerGlyLeuGlyValCysValCysGlySerLe 164
|||||
485 GCAGCTCCATATCTAAGACTGGTGGATTAGGAGATGTTGTGGTCTTT 534
|||||
164 uProIleAlaLeuAlaAlaArgGlyHisArgValMetValMetProA 181
|||||
535 GCCAATGGCAGTACGCTCGGGGTATCGTGAATGGTGGTTCACCTA 584
|||||
181 rGlyrLeuAsnGlySer...SerAspLysAsnTyrAlaLysAlaLeuTyr 196
|||||
585 GGTATTGATGAGGTCTTCAGATGAAAGTACGCCAATGCTGTGAC 634
|||||
197 ThrAlaLysHisIleLysIleProCysPheGlyGlySerHisGluValTh 213
|||||
635 CTTGATGTGGGGCCACTGTCCATTTGTTGGTGTATGCACAGGAGTAGC 684
|||||
213 rPhePheHisGluTyrArgAspAsnValAspTrpValPheValAspHisP 230
|||||
685 CTTCTACCATGAATACAGGCGAGGTGTGATTGGGTATTGTGGACCACT 734
|||||
230 roSerTyrHisArgProGlySerLeuTyrGlyAspAsnPheGlyAlaPhe 246
|||||
735 CTTCTTACTGCAGACCTGGAAACCCCATATGGTGATATTATGGTGCATTT 784
|||||
247 GlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAl 263
|||||
785 GGTGATAAATCAGTTTCGCTTCACTTTGCTTCTCAGCAGCATGTGAAGC 834
|||||
263 aProLeuIleLeuGluGlyTyrIleTyrGlyGlnAsnCysMetP 280
|||||
835 GCCATTGGTCTTCCACTGGGAGGGTTCACCTATGAGAGAGAGAGTGTGT 884
|||||
280 heValValAsnAspTrpHisAlaSerLeuValProValLeuLeuAlaLa 296
|||||
885 TTCTCGCTAATGATTGGCATGCTGCCCTGTTCTTACTTTTAGCGGCC 934
|||||
297 LysTyrArgProTyrGlyValTyrArgSerArgSerThrLeuValI 313
|||||
935 AAGATCGCTCTATGGTGTTTACAAGGATGCTCGTAGTATTGTCGAAT 984

```



```
612 uGlyLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHisAla.. 628
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1885 GGGATTTGATGAGGAGGAGTATGGGAAGGAGGAGTATCTCTGGGAAATGCAG 1934

629 proSerSerThrSerArgSerSerGlyPro 639
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1935 CCATTCAATATGAAAGATTTTCCACCTGGGCGCT 1967

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:US-08-836-567-3

seq_documentation_block:
; Sequence 3, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836.567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1758 base pairs
; TYPE: nucleotide
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Solanum tuberosum
; STRAIN: cv. Berollina
; TISSUE TYPE: tuber tissue
; IMMEDIATE SOURCE:
; LIBRARY: cDNA-library in pBluescriptSKII+
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1377
; OTHER INFORMATION: /function= "Polymerization of
; OTHER INFORMATION: starch"
; OTHER INFORMATION: /product= "Starch synthase"
; US-08-836-567-3
```



```
alignment_scores:
  Quality: 1731.50      Length: 528
  Ratio: 4.142          Gaps: 7
  Percent Similarity: 79.167  Percent Identity: 61.364

alignment_block:
  US-09-674-824-2 x US-08-836-567-3 ..
  Align seg 1/1 to: US-08-836-567-3 from: 1 to: 1758

203 lleProCysPheGlySerHisGluValThrPhePheHisGluTyrAr 219
   ::::::::::::::::::::::::::::::::::::::::::::::::::::
40 GTCCATTGCTTGGTGATGTCAGACAGAGTAGCCTTCTACCATGAATACAG 89

219 gAspAsnValAspTrpValPheValAspHisProSerTyrHisArgPro 236
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
90 GGCAGGTGTTGATGGGTATTTGGACCACCTCTCTTACCGCAGACCTG 139

236 lySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArg 252
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
140 GAAGCCCATATGGTGATATTATGGTGCATTTGGTGAATAATCAGTTCCG 189

253 TyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuLeuLeuLe 269
   ::::::::::::::::::::::::::::::::::::::::::::::::::::
190 TTCATTTCCTTCTCAGCAGCATGTGAAGCGCATTTGGTTCCTCCACT 239

269 uGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTrpH 286
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
240 GGGAGGGTTCATTATGAGAGAGAGTCTGTTCTCGCTAATGATGCA 289

286 lsAlaSerLeuValProValLeuLeuAlaLysTyrArgProTyrGly 302
   ::::::::::::::::::::::::::::::::::::::::::::::::::::
290 ACGTGCCTTGGTTCCTTTATTTTAGCGGCAAGTATCGPCCTTATGGT 339

303 ValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisG 319
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
340 GTTTACAGAGAGTCTGCTAGTATTGTCGCAATACACAACTGTCACATCA 389

319 nGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProGluT 336
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
390 GGGAGTGCAGGCTCGAGTAACCTACAATAATTTGGGTTTGCCTCCACAAT 439

336 rpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAla 352
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
440 GGTATGAGCAGAGTTGAATGGATATTTCCACATGGCAAGGGCGCATGCG 489

353 LeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValThrAl 369
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
490 CTTGCACACTGGTGAACAGTGAACGTTTGAAGGGGCAATAGCAGTTGC 539

369 aAspArgIleValThrValSerGlnGlyTyrSerTrpGluValThrThrA 386
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
540 TGATCGGATACTGACAGTATGCCAGGACTACTCATGGGAATAACAACCTC 589

386 laGluGlyGlyGlnGlyLeuAsnGlnLeuLeuSerSerArgLysSerVal 402
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
590 CTGAAGGGGGATATGGGCTACATGAGCTGTTGAGCAGTAGACAGTCTGTT 639

403 LeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrTh 419
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
640 CTTAATGGAATTAATGAAGTATGATGTTAATGATTGGAACCGCTGCGAC 689

419 rAspLysCysLeuProHisIstYrSerValAspAspLeuSer.....G 434
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
690 AGATGACATATCGCTTCGCATTAATCTCATCAATGACCTCTCCCCCTG 739

434 lyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
740 GAAAGGTTTCAGTCAGACAGTATCTGCAAAAGGAACCTGGGCTTCCAAT 789

451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
   ||| ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```
790 CGACCCGATTGTCCTCACTGATTGGATTATTGGAAGGCTGGACTACAGAA 839
467 sGlyIleAspLeuLysMetAlaIleProGluLeuMetArgGluAspV 484
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
840 AGGTGTTGACATAATCCTGTGAGCAATTCAGAACTTATGCAAGATGATG 889

484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTyrMet 500
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
890 TCCAAAGTTGTAATGCTGGATCTGTGAGAAACAATATGAAGACTGGATG 939

501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
940 AGACATACAGAAATCTTTTAAAGACAAATTTCTGCTGGGTGGATT 989

517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeu 534
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
990 TAATGTTCCAGTTTCTCATAGTAACAGCAGGATGCGACATATCTATGA 1039

534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
1040 TGCCTCAAGATTCAACCGGTGCTTAAACCAATTTGATGCAATGAGA 1089

551 TyrGlyThrValProValIleHisGlyThrGlyGlyLeuArgAspThrVa 567
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
1090 TATGCCACCATACCTATTGTTTCATAGCAGCGGGGSCCTAAGAGACACAGT 1139

567 lGluThrPheAsnProPheGlyAlaLysGlyGlu...GluGlyThrGlyT 583
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
1140 GAAGGATTTTAATCCATATGCTCAAGAAAGAAAGGTGAAGGTACCGGT 1189

583 rpAlaPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThr 599
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
1190 GGACATTTTCTCCTCAACGAGTGAAGGTTGTTTGTATACATCAAGCTG 1239

600 AlaMetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLy 616
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
1240 CGCATCAGGACTTATACAGACATAAGTCATCTTGGGAGGAGTGTGATGAA 1289

616 sArgGlyMetThrLysAspHisThrTrpAspHisAla.ProSerSerThr 632
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
1290 GAGAGGTATGGAGGAGCTATCTCTGGGAAATGTCAGCATTCATCATG 1339

633 SerArgSerSerGlyPro..... 639
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
1340 AGCAAGTTTTCACCTGGGCTTTTATAGATCTCCATATGTCAGATGATT 1389

640 .....SerTrpThrAsnProThrSerCysArgArgGlyL 651
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
1390 ATCAAGAAAGATTGCAAAACGGGATACATCTTAAACTATACGCAAGCTT 1439

651 euGly.....ArgSerLysCysGluSerPro 659
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
1440 TTGGTGTATTAGTACTGTCTATTGGCGCGGAATGTTGT..... 1480

660 SerAlaLeuLysThrSerSerSerPheArgGlyProGluGlyTyrPr 676
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
1481 .....GTTCTTCTGATTTCAGAGAGATCAAGTAGTCTCC 1515

676 oCysThrLeuArgCysProAlaThrVal..... 685
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
1516 AAAGACATGTAGCTGCGCTGTCTGTGTGATGAAGTAAACATAACAAGGCA 1565

686 .....GluSerGlnCysAlaCysLeuLeuTrp 694
   | ::::::::::::::::::::::::::::::::::::::::::::::::::::
1566 ATTAGAAACCCCAACCACTGCTCTCTTTGG 1597

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-941-445A-10
seq_documentation_block:
; Sequence 10, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
```



APPLICANT: Guan, Hanning  
 TITLE OF INVENTION: Starch Encapsulation  
 NUMBER OF SEQUENCES: 37  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
 STREET: 5370 Manhattan Circle  
 CITY: Boulder  
 STATE: CO  
 COUNTRY: US  
 ZIP: 80303  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/941,445A  
 FILING DATE: 30-SEP-1997  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/026,855  
 FILING DATE: 30-SEP-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Winner, Ellen P  
 REGISTRATION NUMBER: 28,547  
 REFERENCE/DOCKET NUMBER: 89-97  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303) 499-8080  
 TELEFAX: (303) 499-8089  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2097 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Zea mays  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..2097  
 US-08-941-445A-10

alignment\_scores:  
 Quality: 1246.00 Length: 714  
 Ratio: 2.800 Gaps: 24  
 Percent Similarity: 62.325 Percent Identity: 42.857  
 alignment\_block:  
 US-09-674-824-2 x US-08-941-445A-10 ..  
 Align seg 1/1 to: US-08-941-445A-10 from: 1 to: 2097  
 7 GlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAspProAlaTh 23  
 |||||  
 65 GCGCAGCGCGG...GCAGTGGGTGCTCTCGCTCGCTACGCGCTAC 111  
 23 rAlaAlaArgAlaSerAlaCysValValArg.....AlaArgLeuArgA 38  
 |||||  
 112 AGCGGCGGCGAGCTCGGTGATTTGGGCGGCGGCGGCGGCGCTC...A 158  
 38 rGLeuAlaArgGlyArgTyrValAlaGluLeuSerArgGluGlyProAla 54  
 |||||  
 159 GGATGAGGCGGCGTGGTACGCGCGCGCAGCGCGCGCGCGGCGGAA 208  
 55 AlaArgProAlaGlnGlnGlnGlnLeuAlaProPro...LeuValPro.... 69  
 |||||  
 209 GCGAGG.....AGGCAGCGCAGAGCTCTCTCTCTCTCTCTCTCTCT 246  
 70 .....GlyPhe..LeuAlaProPr 75  
 |||||

247 GGCCTCTGTTGAGGCGACGACGCGCAAGGCTGTGGATTCTGCTTCACTCC 296  
 75 oProPro.....  
 |||||  
 297 CATCTCTTTCATCTCTCTCGAAGCAAAATCAGAGCGCTGCAATGCAAA 346  
 78 .....AlaProAlaGlnSer 82  
 |||||  
 347 ACGAACGAGTGGGGGCGAGCGCGAGCAGCGCGCGCGCTGTCGCGA 396  
 83 ProAlaProThrGlnProProLeuPro.....AspAl 93  
 |||||  
 397 CCCAAGCTGATCATCATCATCTGCTGTCACCAAGAGAGAAATCTGATGC 446  
 93 aGly.....ValGlyGluLeuAlaProAspLeuLeuL 104  
 |||||  
 447 CAGTGGCGTGAAGCCGAGCGCGAGCTGATGCTAGACCGCTGCGAAA 496  
 104 euGluGlyIleAlaGluAspSerIleAspSerIleIleValAlaAlaSer 120  
 |||||  
 497 GCATAGGATCGCTGAA...CCGGTGGATGCTAAGGCTGATGCGAGCTCCG 543  
 121 GluGlnAspSer.....GluIleMetAspAlaAsnG 131  
 |||||  
 544 GCTACAGATCGCGCGGAGTGTCTTATGACAGAGGAGGATAATGAACC 593  
 131 uGlnProGlnAla...LysValThrArgSerIleValPheValThrGlyG 147  
 |||||  
 594 TGGCCCTTGGCTGGCGCTAATGTGATGAAGCTGCTGCTGGTGGCTTCTG 643  
 147 luAlaAlaProTyrAlaLysSerGlyGlyLeuGlyAspValCysGlySer 163  
 |||||  
 644 AATGTGCTCTCTTCTGCAAGACAGGTGGCTTGGAGATGCTGGTGGTCT 693  
 164 LeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValValMetPr 180  
 |||||  
 694 TTGCTAAGGCTCTGGCGAGGAGGACACCGCTTATGCTGCTGATACCC 743  
 180 oArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrT 197  
 |||||  
 744 AAGATAT.....GGAGAGTATGCGGAGCGCGGATT 775  
 197 hrAlaLysHisIleLysIleProCysPheGlyGlySerHisGluValThr 213  
 |||||  
 776 TAGGTGTAAGGAGAGCTTACAGGTAGCTGGACAGGATTTCAGAGTTACT 825  
 214 PhePheHisGluTyrArgAspAsnValAspTrpValPheValAspHisPr 230  
 |||||  
 826 TATTTTCACTCTTACATTTGATGGATTTGATTTGATTTGATTTGATTTG 875  
 230 oSerTyr...HisArgProGlySerLeuTyrGlyAspAsnPheGlyAlap 246  
 |||||  
 876 TCCCTTCCGCGCAGCGCACATAATATTTATGGG.....GGAGAAA 916  
 246 heGlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAlaCysGlu 262  
 |||||  
 917 GATTCGATATTTTGNAGCGCATGATTTTGTCTGCAAGCGCGCTCTTGAG 966  
 263 AlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGln...AsnCy 278  
 |||||  
 967 GTTCATGATGCTGCTCCATGTTGGCGGTACTGTCTATGTTGATGGCAACT 1016  
 278 smetPheValValAsnAspTrpHisAlaSerLeuValProValLeuLeuA 295  
 |||||  
 1017 AGTTTTCATTTGCTAATGTTGCTATGCGGCTCTCTGCTGCTATCTAA 1066  
 295 laAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLeu 311  
 |||||  
 1067 AGGCTATATACCGGACAAATGTTTGTGATGCGAGTATGCTCGCTCTGCTT 1116  
 312 ValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrPr 328  
 |||||  
 1117 GTGATACACACATTTCTCATCAGGCTGTTGGCCCTGTAGACGACTTCGT 1166



328 oAspLeuGlyLeuProGluTrpTrpTyrGlyAlaLeuGluTrpValPheP 345  
1167 CAATTTTGACTTGCGT..... 1182  
345 roGluTrpAlaArgArgHisAlaLeuAspLys..... 355  
1183 .....GNACACTCATCGACACCTTCAAACTGTATGACAAC 1218  
356 .....GlyGluAlaValAsnPheLeuGlySGLyAlaValThrAlaAs 370  
1219 ATTGGTGGGGATCAGACAACGTTTGTCTGGGGCTGAAGACGCGAGA 1268  
370 pArgIleValThrValSerGlnGlyTyrSerTrpGluValThrAlaG 387  
1269 CCGGCTGGTGACCGTTAGCAATGGCTACATGTGGAGCTCAAGACTTCG 1318  
387 luGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerValLeu 403  
1319 AAGCGCGTGGGGCCTCCACGACATCATAAACACGACACCTGGAAGCTG 1368  
404 AsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAs 420  
1369 CAGGCATCGTGAACGGCATCGACATGACGAGTAGGAAACCCGCTGTGA 1418  
420 pLysCysLeuProHis.....HisTrpSerValAspAspLeu. 432  
1419 CGTGCACCTCCACTCCGACGACTACACCACTACACCTTCGAGACGCTG 1468  
433 ..SerGlyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeu 448  
1469 ACACCGCAACGGCAGTGCAGGCGGCCCTGCACGCGCAGCTGGGCTG 1518  
449 ProValArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTr 465  
1519 CAGTTCGCGAGCAGCTGCCACTGATCGGGTTCATCGGCGCGCTGCACCA 1568  
465 rGlnLysGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArg 482  
1569 CCAGAAAGGCGGTGCACATCATCGCGACGCGATCCACTGGATCGCGGGC 1618  
482 luAspValGlnPheValMetLeuGlySerGlyAspProIlePheGluGly 498  
1619 AGGACGTGCAGCTCGTATGCTGGCACCGCGGCGCGCGCTGGAGAC 1668  
499 TrpMetArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpVa 515  
1669 ATGCTCGCGGGTTCGATCGGACACGACCAAGTTCGCGGTGGGT 1718  
515 lGlyPheSerValProValSerHisArgIleThrAlaGlyCysAspIleL 532  
1719 GGGGTTCGCTGGTGCCTTCGCGCACCGCATCACGCGGCGCGGACATCC 1768  
532 euLeuMetProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAla 548  
1769 TGTGTATGCGCTCGCGGTTAGACCGTTCGGGCTGAACCACTCTAGGCC 1818  
549 MetGlnTyrGlyThrValProValIleHisGlyThrGlyGlyLeuArgAs 565  
1819 ATGGCGCTACGGGACCGTGCCTGGTGCACGCGTGGGGGGCTCCGGGA 1868  
565 pThrValGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrG 582  
1869 CACGGTGGCGCCGTTGACCCGTT.....AAGCAGCACCGCGGCTCG 1909  
582 lyTrpAlaPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArg 598  
1910 GGTGGACGTTTCACCGCGAGCGAACCAGGATGATCGACGCGCTCTCG 1959  
599 ThrAlaMetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMe 615  
1960 CACTGCCTCACCACGTACCGAACTACAGAGAGAGTGGCGCGCTCGAC 2009

```

615 LysArgGlyMetThrLysAspHisThrTrpAspHisAla.Pro..... 629
|||||.....||| :|||||.....|||
2010 GGCGCGGATGCCGAGGACCTCACTGGGACCACGCGCGCTGCTGT 2059
630 .....SerSerThrSerArgSerSerGly 638
|||||.....|||
2060 ATGAGGAGCGTCTCTCAAGCGCAGTAGTACCAGTGT 2095

seq_name: /cgn_2_6/ptodata/1/1/ina/6A_COMB.seq:US-08-836-567-7

seq_documentation_block:
; Sequence 7, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2793 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Solanum tuberosum
; STRAIN: cv Désire
; TISSUE TYPE: leaf tissue
; IMMEDIATE SOURCE:
; LIBRARY: cDNA-library in Lambda ZAP11
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 242..2542
; US-08-836-567-7

```

alignment_scores:		
Quality:	1215.00	Length: 578
Ratio:	3.030	Gaps: 12
Percent Similarity:	69.377	Percent Identity: 45.675



alignment block:

US-09-674-824-2 x US-08-836-567-7

Alian seq 1/1 to: US-08-836-567-7 from: 1 to: 2793

79 ProAlaGlnSerProAlaProThrGlnProProLeu..... 90  
||| : : : : : ||| : : : : : ||| : : : : :  
824 CTTCCAGAAACTCCGAAGTCCAGTCAAGAGACACATTTGGATGTGAATTC 873  
91 : : : : : .....ProAspAlaGlyValGlyGluLeuAlaLap 100  
||| : : : : : ||| : : : : :  
874 ACCCAAAAGTTTACTAGATGTTCTCGGAAGAAGATCCAGTCTTATATGC 923  
100 roAspLeuLeuGluGlyLeuAla.....GluAsp 110  
||| : : : : : ||| : : : : : ||| : : : : :  
924 CTTTCATTACGTAAGAAGTCTCAGCAGTATSerGlnAspSerGluLeu 127  
111 SerLeuAspSerLeuLeuValAlaLaserGlnAspSerGluLeu 127  
: : : : : ||| : : : : : ||| : : : : :  
974 ATCTTTGAAGATCAAGTCTGAGGCAACAGAGAGACTGAAGATCCTGT 1023  
127 tAspAlaAsnGluInProGlnAlaLysVal.....ThrArgSerI 141  
: : : : : ||| : : : : : ||| : : : : :  
1024 GAATATAGATGAGAAACCCCTCCATTTGCAGCAACAAATGTTATGAACA 1073  
141 leValPheValThrGlyGluAlaAlaProTyTAlaLysSerGlyGlyLeu 157  
: : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
1074 TTATTTGGTGGCTTCAGATGGCTCCATGGTCTAAACAGGTGGGCTT 1123  
158 GlyAspValCysGlySerLeuProLeuAlaLeuAlaAlaArgGlyHisAr 174  
: : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
11124 GGAGATGTTGCTGGAGCATTTACCCAAGCTTTGGCTGCAGCGTGGCCACAG 1173  
174 gValMetValValMetProArgTyTLeuAsnGlySerSerAspLysAsnT 191  
: : : : : ||| : : : : : ||| : : : : :  
1174 AGTTATGGTGTGGCCTCGTTATGACAACTATCTGAACCTCAAGATT 1223  
191 yrAlaLysAlaLeuTyTThrAlaLysHisLysLysLeuProCysPheGly 207  
: : : : : ||| : : : : : ||| : : : : :  
1224 CTGGT.....GTAGAAAAATTTATAAGTTGAT.....GGT 1255  
208 GlySerHisGluValThrPhePheHisGluTyTArgAspAsnValAspTr 224  
: : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
1256 CAGGATGGGAAGTGACTACTTCCAAGCTTTTATTGATGGTGTGGATT 1305  
224 pValPheValAsp...HisProSerTyTHisArgProGlySerLeuTyTArg 240  
: : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
1306 TGTTTTCATTGACAGTCATATGTTTAGACACATCTGGGAACAACATTTACG 1355  
240 LyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyTThrLeu 256  
: : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
1356 GAGGGAAC.....CGTGTGGATATTTTAAACGCATGGTTTATTT 1396  
257 CystTyTAlaAlaCysGluAlaProLeuLeuGluLeuGlyTyTII 273  
||| ||||| ||||| ||||| : : : : : ||||| : : : : :  
1397 TGCAAGACGCGATTGGGTTCCCTGGCATGTTCCATGTGGTGGGCTG 1446  
273 eTyT GlyL...AsnCysMetPheValAlaAsnAspTrpHisAlaSerL 289  
: : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
1447 CTATGGAGATGGAATTTAGTGTTCATTGCTATCATGTTGCCATCTGCTT 1496  
289 euValProValLeuLeuAlaLysTyTArgProTyTyrGlyValTyTArg 305  
||| ||||| ||| ||| ||||| : : : : : ||| : : : : :  
1497 TATTCGCCAGTATATCTGAAAGCTTATTATCTGCACAATGGAATTTATGAAC 1546  
306 AspSerArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGl 322  
: : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
1547 TATACAAGATCTGCTCGGTGATTCATATCAATCCCTCATCAGGTCGTGG 1596  
322 uProAlaSerThrTyTProAspLeuGlyLeuProProGluInTyTyrGlyA 339  
: : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
1597 TCCTTTTGAGGATTTTTCATATGTAGATGTTCACACCATATATGACAC 1646

339 laLeuGIutrpValPheProGIuTrpAlaArGArHisAlaLeuAspLys 355  
 : : : : : |||  
 1647 CTTTCAAGTTGTATGACCCAGTA. : : : : : . . . . . GGA 1672

356 GlyCluAlaValAlaSnPheLeuLysGlyAlaValValThrAlaAspArgIle 372  
 ||||| : : : : : : : : : : : ||||| : : : : :  
 1673 GGTGAGCATTTCAACATTTTGGCGCTGGTCTAAAGACACAGATCGTGT 1722

372 eValThrValSerGlnGlyTyrSerTrpGluValThrAlaGluGlyG 389  
 : : : : : ||||| : : : : : : : : : : : ||||| : : : : :  
 1723 AGTTACAGTTAGTCATGGATATTTCATGGGAACCTAAAGACTTCCCAAGGTG 1772

389 lyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGly 405  
 || :  
 1773 GTTGGGATTTGCATTCAGATTAATTAAATGACACCGATTGGAAATTTACAGGCT 1822

406 IleValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCys 422  
 ||||| : : : : : : : : : : : ||||| : : : : :  
 1823 ATTGTGAATGGGATTGATACAAAGAGTGGAAACCTTGAGTTGACCTTCA 1872

422 sLeuProHis. : : : : : . HisTyrSerValAspAspLeu. SerG 434  
 ||||| : : : : : : : : : : : ||||| : : : : :  
 1873 CTTACAGTCAGATGTTTACATGAACACTCTCTTGGACAGCTACAGACTG 1922

434 LyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450  
 : : : : : : : : : : : ||||| : : : : : : : : : :  
 1923 GCAGGCTCAATGATTAAGCTTCGATTCGAGAAGAGACTTGGTTTACAGTT 1972

451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnIy 467  
 : : : : : : : : : : : ||||| : : : : : : : : : :  
 1973 CGTGATGATGTCACCATGATCGGTTTCATTTGGGAGGCTTGACCCACAAA 2022

467 sClyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484  
 ||||| : : : : : : : : : : : ||||| : : : : :  
 2023 GGGTGTTCATCTGATGCTGAGGCCAGCTGTTGGATGATGGGTACGAGTNG 2072

484 aIcInPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500  
 : : : : : : : : : : : ||||| : : : : : : : : : :  
 2073 TACAACCTGGTCATGTTGGGAGCGGGAGCTGACCTTGAACAGATGCTA 2122

501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517  
 : : : : : ||||| : : : : : : : : : : : ||||| : : : : :  
 2123 AGGCAATTTGATGTGTCACACAAATGATAAAATAGAGGATGGGTGGTTT 2172

517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534  
 : : : : : : : : : : : ||||| : : : : : : : : : :  
 2173 CTCTGTGAAGACTTCTCATGCTATAACTGCTGGTGCAGACATTTCTGCTCA 2222

534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550  
 : : : : : : : : : : : ||||| : : : : : : : : : :  
 2223 TGCCITCTAGATTTCAGCCTTGGGACTGGAACACAGCTTTATGCAATGAAA 2272

551 TyrGlyThrValProValIaHisClyThrGlyGlyLeuArgaspThrVa 567  
 : : : : : : : : : : : ||||| : : : : : : : : : :  
 2273 TATGGGACTATTTCGTGTGTCATGCAATGAGGAGGACTCAGAGATACTGT 2322

567 lGIuThrPheAsnProPheGlyAlaLysGlyCluGlyThrGlyTrpA 584  
 : : : : : : : : : : : ||||| : : : : : : : : : :  
 2323 GCAGCCCTTTGATCCTTT. : : : : : . AATGAGTCAGGACTGGGCTGGA 2363

584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600  
 : : : : : : : : : : : ||||| : : : : : : : : : :  
 2364 CCTTCAGTAGGGCTGAAGCTACCGAGCTGATCCAGCGCAATTAGGAATATGC 2413

601 MetSerThrPheArgCluHisIysProSerTrpGluGlyLeuMetLysAr 617  
 : : : : : : : : : : : ||||| : : : : : : : : : :  
 2414 TTACTGACTTATCTGTGTAGTACAAAAGAGTGGGAGGGGATTCAGACACG 2463

617 gGlyMetThrLysAspHisThrTrpAspHisAla 628  
 : : : : : : : : : : : ||||| : : : : : : : : : :  
 2464 TTCTATGACACAGACTTAAAGTTGGGATTAATGCT 2497

2464 TTGTATGCACACAAGACTTAAGTTGGGATAATGCT 2497



seq\_name: /cgn2\_6/ptodata/1/1na/6A\_COMB.seq:US-08-836-567-5

```

seq_documentation_block:
; Sequence 5, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1926 base pairs
; TYPE: nucleotide
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOPHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Solanum tuberosum
; STRAIN: cv. Berollina
; TISSUE TYPE: tuber tissue
; IMMEDIATE SOURCE:
; LIBRARY: cDNA-library in pBluescriptSK+
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1675
; OTHER INFORMATION: /function= "Polymerization of
; OTHER INFORMATION: starch"
; OTHER INFORMATION: /product= "Starch synthase"
US-08-836-567-5

```

alignment\_scores:  
 Quality: 1184.00 Length: 555  
 Ratio: 3.020 Gaps: 11  
 Percent Similarity: 70.631 Percent Identity: 46.306

alignment\_block:

US-09-674-824-2 x US-08-836-567-5 ..

Align seg 1/1 to: US-08-836-567-5 from: 1 to: 1926

```

90 LeuProAspAlaGlyValGlyLeuAlaProAspLeuLeuGluGlu 106
   ::::: ::::: ::::: ::::: ::::: ::::: :::::
26 GTTCTGGAAGAGATCCAGTCCTATATGCTTCATTACGTAAGAATC 75
   ::::: ::::: ::::: ::::: ::::: :::::
106 yllsala.....GluAspSerIleAspSerIleLeu 117
   : ||| ::::: ::::: ::::: :::::
76 CTCAGCATCCCATGTGGAACAGAGAAATGAAATCTTGAAGGATCAAGTG 125
   ::::: ::::: ::::: ::::: ::::: :::::
117 alAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnPro 133
   ::::: ::::: ::::: ::::: ::::: :::::
126 CTGAGGCAACAGAGAGACTGAAGATCTGTGAATATAGATGAGAAACCC 175
   ::::: ::::: ::::: ::::: ::::: :::::
134 GlnAlaLysVal.....ThrArgSerIleValPheValThrGlyGlu 147
   ::::: ::::: ::::: ::::: ::::: :::::
176 CCTCCATTGGCAGGAACAATGTTATGAACATTATTTTGGTGGCTTCAGA 225
   ::::: ::::: ::::: ::::: ::::: :::::
147 uAlaAlaProTyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerL 164
   ::::: ::::: ::::: ::::: ::::: :::::
226 ATGCGCTCCATGGTCTAAACAGCTGGCTGGAGATGTTGCTGGAGCAT 275
   ::::: ::::: ::::: ::::: ::::: :::::
164 euProIleAlaLeuAlaAlaArgGlyHisArgValMetValValMetPro 180
   ::::: ::::: ::::: ::::: ::::: :::::
276 TACCAAGCTTGGCTCGAGCTGGCCACAGATTATGTTGTGGACCT 325
   ::::: ::::: ::::: ::::: ::::: :::::
181 ArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrTh 197
   ::::: ::::: ::::: ::::: ::::: :::::
326 CGTTATGACAACATATCTCTGAACCTCAAGATTCTGGT.....GT 363
   ::::: ::::: ::::: ::::: ::::: :::::
197 rAlaLysHisIleLysIleProCysPheGlyGlySerHisGluValThrP 214
   ::::: ::::: ::::: ::::: ::::: :::::
364 AAAAAAATTATAAGTTGAT.....GGTCAGGATGTGGAAGTGACTT 407
   ::::: ::::: ::::: ::::: ::::: :::::
214 hePheHisGluTyrArgAspAsnValAspTyrPheValAsp...His 229
   ::::: ::::: ::::: ::::: ::::: :::::
408 ACTTCAAGCTTTTATTGATGGTGGATTTTGTTCATTGACAGTCAT 457
   ::::: ::::: ::::: ::::: ::::: :::::
230 ProSerTyrHisArgProGlySerLeuTyrGlyAspAsnPheGlyAlaPh 246
   ::::: ::::: ::::: ::::: ::::: :::::
458 ATGTTTAGACACATATGGGAACAACATTTACGGAGGAAC.....CG 498
   ::::: ::::: ::::: ::::: ::::: :::::
246 eGlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAlaCysGluA 263
   ::::: ::::: ::::: ::::: ::::: :::::
499 TGTGGATATTTTAAACGCATGTTTATTTTTCAAAGCAGCGATTGAGG 548
   ::::: ::::: ::::: ::::: ::::: :::::
263 laProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGln...AsnCys 278
   ::::: ::::: ::::: ::::: ::::: :::::
549 TTCCTTGGCATGTTCATGTGGTGGGTCTGCTATGGAGATGGAATTTA 598
   ::::: ::::: ::::: ::::: ::::: :::::
279 MetPheValValAsnAspTyrPheHisAlaSerLeuValProValLeuAl 295
   ::::: ::::: ::::: ::::: ::::: :::::
599 GTGTTTCAATGCTAATGATGGCATCTGCTTTTATGCCAGTATATCTGAA 648
   ::::: ::::: ::::: ::::: ::::: :::::
295 aAlaLysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLeuV 312
   ::::: ::::: ::::: ::::: ::::: :::::
649 AGCTTATTATCGTGACAATGGAATTATGAATAACAAGATCTGTCTGG 698
   ::::: ::::: ::::: ::::: ::::: :::::
312 alIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrPro 328
   ::::: ::::: ::::: ::::: ::::: :::::
699 TGATTCATAACATCGCTCATCGGTCTGCTCTTTGGAGGATTTTCA 748
   ::::: ::::: ::::: ::::: ::::: :::::
329 AspLeuGlyLeuProGluThrTyrGlyAlaLeuGluThrValPhePr 345
   ::::: ::::: ::::: ::::: ::::: :::::
749 TATGTAGATCTTCCACCACACTATATGGACCCCTTTCAGTGTGATGACCC 798
   ::::: ::::: ::::: ::::: ::::: :::::
345 oGluTrpAlaArgHisAlaLeuAspLysGlyGluAlaValAsnPhel 362
   ::::: ::::: ::::: ::::: ::::: :::::
799 AGTA.....GGAGGTGAGCATTTCAACATTT 824
   ::::: ::::: ::::: ::::: ::::: :::::
362 euLysGlyAlaValValThrAlaAspArgIleValThrValSerGlnGly 378
   ::::: ::::: ::::: ::::: ::::: :::::

```



825 TTGGCGTGGTCTAAAGACAGCAGATCGTAGTTACAGTTAGTCATGGA 874  
379 TyrSerTrpGluValThrAlaGluGlyClnGlyLeuasnGluLe 395  
875 TATTTCATGGGAACATAAGACTTCCCAAGGTGGTGGGATGTCATGAT 924  
395 uLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAspI 412  
925 AATTAATGAGAACGATTGGAAATTAAGAGGTATTGTGAATGGGATGATA 974  
412 leAsnAspTrpAsnProThrThrAspLysCysLeuProHis..... 425  
975 CAAAAGAGTGAACCCCTGAGTTGACGCTTCACTTACAGTCAGATGTTAC 1024  
426 ...HisTyrSerValAspAspLeu...SerGlyLysAlaLysCysLysAl 440  
1025 ATGAACACTACTCTTGGACACCGCTGAGTTGACGCTTCACTTCAATGTAAGC 1074  
440 aGluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuI 457  
1075 TGCATTGCGAAGGAACTGGTTTACCAGTTCGTGATGATGCCACTGA 1124  
457 leGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLys 473  
1125 TCGGTTTCATTGGGAGGCTTGCACCAAAAAGGTTGTGATCTGATTGCT 1174  
474 MetAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuG 490  
1175 GAGGCCACTGCTTGGATGATGGTCAGGATGTACAACCTGGTCATGTTGGG 1224  
490 ySerGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSert 507  
1225 GACGGGAGGCGTGACCTTGAACAGATGCTAAGGCAATTTCAAGTGTCAAC 1274  
507 yLysAspLysPheArgGlyTrpValGlyPheSerValProValSerHis 523  
1275 ACAATGATAAAATTAGAGGATGGTGGTTGCTTCTGTGAAGACTTCTCAT 1324  
524 ArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluPr 540  
1325 CGTAACTGCTGGTGGACAGACATTCGTCATCGCTTCTAGATTGAGGC 1374  
540 ocysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValV 557  
1375 CTTCGCACTGAACCAAGCTTATGCAATGAAATATGGGACTATTCTGTTG 1424  
557 alHisGlyThrGlyLeuArgAspThrValGluThrPheAsnProPhe 573  
1425 TTCAATGCAGTAGGAGACTCAGAGATATGTCGACCCCTTTGATCCTTTT 1474  
574 GlyAlaLysGlyGluGlyThrGlyTrpAlaPheSerProLeuThrVa 590  
1475 .....AATGAGTCAGGACTGGGTGGGACCTTCAGTAGGCTGAAGC 1515  
590 lAspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluH 607  
1516 TAGCCAGCTGATCCAGCAATAGGAATTTGCTTACTGACTTATCGTGAGT 1565  
607 isLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHis 623  
1566 ACAAAAAGAGTTGGGAGGATTCAGACACGTTGTGTATGACACAGACTTA 1615  
624 ThrTrpAspHisAla 628  
1616 AGTTGGGATAATGCT 1630

seq\_name: /cgn2\_6/ptodata/1/1na/6B\_COMB.seq:US-09-196-390-5  
seq\_documentation\_block:  
; Sequence 5, Application US/09196390  
; Patent No. 6307125  
; GENERAL INFORMATION:  
; APPLICANT: Block, Martina  
alignment\_scores:  
Quality: 1182.50 Length: 825  
Ratio: 2.571 Gaps: 24  
Percent Similarity: 55.758 Percent Identity: 36.727  
alignment\_block:  
US-09-674-824-2 x US-09-196-390-5 ..  
Align seg 1/1 to: US-09-196-390-5 from: 1 to: 2825

APPLICANT: Lorz, Horst  
APPLICANT: Lutticke, Stephanie  
APPLICANT: Walter, Lennart  
APPLICANT: Froberg, Claus  
APPLICANT: Kossmann, Jens  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES  
TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH  
SYNTHESIS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/196,390  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 196 21 588.9  
FILING DATE: 29-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 196 36 917.7  
FILING DATE: 11-SEP-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP97/02793  
FILING DATE: 28-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley, Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: AGREVO-9  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2825 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Triticum aestivum L.  
STRAIN: cv. Florida  
TISSUE TYPE: ca. 21 d Caryopses  
IMMEDIATE SOURCE:  
LIBRARY: cDNA library in pBluescript sk (-)  
CLONE: pTAS1  
NAME/KEY: CDS  
LOCATION: 162..2559  
US-09-196-390-5







```
2004 CTGCTCGCTTTCATCGCGCGCTGGAGCGGAGAAAGCGGTGGAGATCAT 2053
472 eLysMetAlaIleProGluLeuMetArgGluAspValGlnPheValMetL 489
2054 CGCGGAGCCCATCGCTGGATCGTGAGACGAGCGTGCAGCTGGTGATGC 2103
489 euGlySerGlyAspProIlePheGluGlyTrpMetArgSerThrGluSer 505
2104 TGGGACCGGGCCAGCAGCTGGAGAGCATGCTGCAGCACTTCGAGCGG 2153
506 SerTyrLysAspLysPheArgGlyTrpValGlyPheSerValProValSe 522
2154 GAGCACCAGCAAGAGTGGCGGGTGGGTTCCTCCGTGGCGCTGGC 2203
522 rHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheG 539
2204 GCACCGGATCAGCGGGGGGGAGCGGCTCTCATGCCCTCCCGGTTCG 2253
539 luProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValPro 555
2254 AGCGGTGGGGCTGAACAGCTCTACGCCATGGCTACGGCAGCGTCCCG 2303
556 ValValHisGlyThrGlyLeuArgAspThrValGluThrPheAsnPr 572
2304 GTCGTGCACGCGCTCGCGGGCTCAGGAGACACCGTGGCGGCTTCGACC 2353
572 oPheGlyAlaLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuT 589
2354 CTTC.....ACCCTCCGGCTCGGTGGAGCTTCACCGCGCGCG 2394
589 hrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArg 605
2395 AGCGCACAAAGTGATCGAGCGCTCGGGCACTGCCTCCGCACTACCGA 2444
606 GluHisLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAS 622
2445 GACTTCAAGGAGAGCTGGAGGGCCCTCCAGAGCGCGCATGTGCGCAGGA 2494
622 phiThrTrpAspHisAla.ProSerThrSerArgSerSerGly 638
2495 CTTACGCTGGGAGCAGCGCGCCCAAGCTCTACGAGGACGCTCCTCGTCAAG 2544
639 ProSerTrpThrAsn.....ProThrSe 646
2545 CCAAGTACAGTGGTGAACGCTAGCTGCTAGCGCTCCAGCGCCCGCATGC 2594
646 cYsArgArgGlyLeuGly.....ArgSerLysCysGlu...S 658
2595 GTGCATGACAGGTGAAGTGCATTCGCGCAGCAGCAAGTGCATGGAG 2644
658 exProSerAlaLeuLysThrSer 665
2645 CGCCGGCATCCCGGAAGTACAGT 2667
```

seq\_name: /cgm2\_6/ptodata/1/ina/6A\_COMB.seq.US-08-941-445A-8

seq\_documentation\_block:  
Sequence 8, Application US/08941445A  
Patent No. 6107060  
GENERAL INFORMATION:  
APPLICANT: Keeling, Peter  
APPLICANT: Guan, Hanping  
TITLE OF INVENTION: Starch Encapsulation  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GreenLee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/941,445A  
FILING DATE: 30-SEP-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,855  
FILING DATE: 30-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Winner, Ellen P  
REGISTRATION NUMBER: 28,547  
REFERENCE/DOCKET NUMBER: 89-97  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2007 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2007  
US-08-941-445A-8
```

alignment\_scores:  
Quality: 1076.00 Length: 693  
Ratio: 2.485 Gaps: 18  
Percent Similarity: 62.482 Percent Identity: 37.951

alignment\_block:  
US-09-674-824-2 x US-08-941-445A-8 ..

Align seg 1/1 to: US-08-941-445A-8 from: 1 to: 2007

```
5 GlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAspPr 21
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
12 GCCGGGGGGAAGGAGCGCGCGGAGAGGAGCGGCGAGCGCCCGCAG.. 59
21 oAlaThrAlaAlaArgAlaSerAlaCys.ValValArgAlaArgLeuArg 37
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 ....GTTGCCCGCGCTCGGGCAATGGGTCTCCAAACGAGGAGGATCCT 105
38 ArgLeuAlaArgGlyArgTyr.....ValAlaGluLeuSerArgG1 51
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
106 CTTCAGCGCGTCCGCGCGTACGGCTCCGCGAGCGGGAACACGCGCAGGAC 155
51 uGlyProAlaAlaArgProAlaGlnGlnLeuAlaPro..... 65
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 CGGCGCGCGTCC.....TCCAGACGCGCGCATTCGCGGACGTTGAGA 199
66 ..ProLeuValProGlyPheLeuAlaProProProAlaProAlaGln 81
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
200 TCGTTGAGATCAAGTCCATCGTCGCGCGCGCGCAGCAGCATAGTGAAG 249
82 SerProAlaProThrGlnProLeuProAspAlaGlyValGlyGluLe 98
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
250 TTCCCAAGG.....CGCGGGCTACAGATGATCTTCCCTCTGGGACAT 293
98 uAlaProAspLeuLeu..... 104
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
294 AGCAGCGAGAGTGTCTCTCCAGCGCCCGAGCCACTGCATGAATCGCGTG 343
105 .....GluGlyIleAlaGluAspSerIleAspSerIle 115
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



[illegible]

388 GlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAs 404  
1243 GCGGGCTGGGGCTCCACGACATCATCGTTCTTAACGACTCGAAGACAA 1252  
404 nGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspL 421  
1293 TGCATTCGTGAACGATCGACACACGAGGTGGAACCCCAAGGTGGACG 1342  
421 yscysLeuProHis.....HisTrpSerValAspLeu... 432  
1343 TGCACCTGGGGTCGGACGGCTACACCACTACTCCCTCGAGACACTCGAC 1392  
433 SerGlyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuPr 449  
1393 GCTGAAAGCGCGCATGGAAGCGGCCCTCGACGCGGACGTTGGCCCTGGA 1442  
449 oValArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTrpG 466  
1443 AGTGGCGACGACGTGGCGCTGCTGGGCTTCATCGGGCGTCTGGATGGAC 1492  
466 InLysGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGlu 482  
1493 AGAAGGCGTGGACATCATCGGGGACGCGATCGCTGGATCGCGGGCGAC 1542  
483 AspValGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTr 499  
1543 GACGTGCAGCTGGTATGCTGGGACCGGCCACCTCACCTGGAACGAAT 1592  
499 pMetArgSerThrGluSerSerTrpLysAspLysPheArgGlyTrpValG 516  
1593 GCTGCAGACACTGGAGCGGGAGCATCCCAACAAAGTCGGCGGGTGGGTG 1642  
516 lypheSerValProValSerHisArgIleThrAlaGlyCysAspIleLeu 532  
1643 GTTCTCGTCTTAATGTTGTCATCGCATCGCCGCGCGGCCGCGTGTG 1692  
533 LeuMetProSerArgPheGluProCysGlyLeuAsnGlnLeuTrpAlaMe 549  
1693 GTGATGGCCCTCCGGCTTCGCGCGC...GGGCTGAACACAGCTACGCGAT 1739  
549 tGlnTrpGlyThrValProValValHisGlyThrGlyGlyLeuArgAspT 566  
1740 GGCATACGGCACCGTCCCTGTGTGTGCACGCGTGGCGGGGCTCAGGGACA 1789  
566 hrValIgluThrPheAsnProPheGlyAlaLysGlyGluGlyIleThrGly 582  
1790 CCGTGGCGCCGTTCTGACCCGTTT.....GGCGAGCGGGCTCGGG 1830  
583 TrpAlaPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgTh 599  
1831 TGGACTTTTACCGCGCGCGAGGCCAACAGTGTATCGAGTGTCTCAGCCA 1880  
599 rAlaMetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetL 616  
1881 CTGCTCGACACGTACCGAAACTACGAGGAGAGCTGGAAGAGTCTCCAGG 1930  
616 ysArgGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerTh 632  
1931 CGCGGGCATGTCGACAGACCTCAGCTGGGACCAACGCGGCTGAGCTCTAC 1980  
632 rSerArgSerSerGlyProSer 640  
1981 GAGGACGCTCTTGTCAAGTACCACT 2005

seq\_name: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq;US-08-572-951-2

seq\_documentation\_block:

; Sequence 2, Application US/08572951

; Patent No. 5824790

; GENERAL INFORMATION:

APPLICANT: KEELING, PETER L

; APPLICANT: KNIGHT, MARY E.







```
seq_documentation_block:
; Sequence 3, Application
; Patent No. 5824790
; GENERAL INFORMATION:
```

```

78 AlaProAlaGlnSerProAlaProThrGlnProProLeuPro..... 91
   |||||:::   ::::|::|   ::::|::|   |||
439 GCNCCNGTNSWGNCNNAARGCAYCAYCCNWSNCCNCCNTNACNAA 488
   |||||:::   ::::|::|   ::::|::|   |||
92 .....AspAlaGly.....ValGlyGluLeuA 99
   |||||:::   ::::|::|   ::::|::|   |
489 RMGNGARATHGAYGCNWSNCCNGTNAARCCNAGCCNCGNGCAYGATG 538
   |||||:::   ::::|::|   ::::|::|   |
99 laProAspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIle 115
   |||||:::   ::::|::|   ::::|::|   |||
539 CNMGCCNCTNGARWSNATHGGNATHGCGNAR...CCNGTNGAYCGCNAAR 585
   |||||:::   ::::|::|   ::::|::|   |||
116 IleValAlaAlaSerGluGlnAspSer.....GluI 126
   |||||:::   ::::|::|   ::::|::|   |

```



```
586 GCNGAYGCNGCCNGCCNACNGAYGCNGCCNMGNGCCNCCNTAYAYMG 635
126 eMeTAspAlaAsnGluInProGlnAla...LysValThrArgSerIleVal 142
636 NGARGAYAAAYGARCNGCCNGCCNTYNGCCNGCCNCCNTAYATGATGAYG 685
142 alpheValThrGlyCylAlaAlaProTyrAlaLysSerGlyGlyLeuGly 158
686 TNGTNGTNGCCNMGNGCCNTYNGCCNCCNTYTGAAACACNGCCNGGNTNG 735
159 AspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgVa 175
736 GAYGTNGTNGCCNMGNGCCNTYNGCCNCCNTYNGCCNMGNGCCNMGNGT 785
175 lMeTValValMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrA 192
786 NATGTNGTNGCCNMGNTAY...GGNGARTAYG 817
192 laLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheGlyCly 208
818 CNGARGCCNMGNGAYTNGCCNMGNTNGCCNMGNTAYAAARGTNGCCNGCAR 867
209 SerHisGluValThrPhePheHisGluTyrArgAspAsnValAspTrpVa 225
868 GAYWSNGARGTNGCCNMGNTAYTNGCCNMGNTAYTNGCCNMGNTAYTNGT 917
225 lPheValAspHisProSerTyr...HisArgProGlySerLeuTyrGlyA 241
918 NTYGTNGCCNMGNTAYTNGCCNMGNTAYTNGCCNMGNTAYTNGCCNMGNT 966
241 spAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCys 257
967 .....GGNGARMGNTNGAYATHTYNAARMGNTATGATHTYNTTYG 1008
258 TyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTy 274
1009 AARGCCNMGNTNGCCNMGNTAYTNGCCNMGNTAYTNGCCNMGNTAYTNG 1058
274 rGlyGln...AsnCysMetPheValValAsnAspTrpHisAlaSerLeuV 290
1059 YGNGAYGCAAYTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGT 1108
290 alProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAsp 306
1109 TWCCNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 1158
307 SerArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluPr 323
1159 GCNMGNSNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 1208
323 oAlaSerThrTyrProAspLeuGlyLeuProProGluTrpTyrGlyAlaL 340
1209 NGTNGAYGAYTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGT 1239
340 euGluTrpValPheProGluTrpAlaArgArgHisAlaLeuAspLys... 355
1240 .....GARGATATATGAYCAYTYT 1260
356 .....GlyGluAlaValAsnPheLeuLysGlyAl 365
1261 AARYTNTAYGAYAAAYATHGGNGNGAYCAYWSNAAAYGTNTTYGCGCNGG 1310
365 aValValThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpG 382
1311 NYTNAARACNGCCNGAYMGNTNGTNGCCNMGNTNGTNGCCNMGNTNGT 1360
382 luValThrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuSerSer 398
1361 ARYTNAARACNGCCNMGNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 1410
399 ArgLysSerValLeuAsnGlyIleValAsnGlnGlyIleAspIleAsnAsp 415
1411 AAYGAYTGCAARYTNCARGCNATHGNTNGTNGTNGTNGTNGTNGTNGTNG 1460
```

```
415 pAsnProThrThrAspLysCysLeuProHisHisTyrSerValAspAspL 432
1461 GAAYCCNGCCNGTNGAYGTN.....CAYTNCAYWSNGAYGAYT 1498
432 eu.....SerGlyLysAlaLysCysLys 439
1499 AYACNAAAYTAYACNTTYGACRACNTYNGAYACNGGNAARMGNCARTGYAAR 1548
440 AlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLe 456
1549 GCNCNTNCCARMGNCARTYNGGNTNCCARTNMGNGAYGAYGTNCCNYT 1598
456 uileGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleL 473
1599 NATHGGNTTYATHGGNMGNTNGAYCAYCAARAARGGNGTNGAYATHATHG 1648
473 ysMetAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeu 489
1649 CNGAYGCNATHCAYTGGATHGCGGNCARGAYGTNCARTYNGTNGTNGTNG 1698
490 GlySerGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSe 506
1699 GGNACNGGNGCCNGAYTNGARGAYATGYTNMNGMNTTYGARWSNGA 1748
506 rTyrLysAspLysPheArgGlyTrpValGlyPheSerValProValSerH 523
1749 RCAYWSNGAYAAARGTNGCCNMGNTGGTNGGNTTYWSNGTNGCCNMGNC 1798
523 isArgIleThrAlaGlyCysIleLeuLeuMetProSerArgPheGlu 539
1799 AYMGNATHACNGCCNGCCNGAYATHTYNTNATGCCNWSNMGNTTYGAR 1848
540 ProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProVa 556
1849 CCNTGYGNTNAAAYCARYTNTAYGCNATGCCNTAYGNGACNGTNGCCNGT 1898
556 lValHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProp 573
1899 NGTNCAYGCNTNGGNGGNTNGMNGAYACNGTNGCCNMGNTTYGAYCCT 1948
573 heGlyAlaLysGlyGluGlyThrGlyTrpAlaPheSerProLeuThr 589
1949 TY.....AAYGAYACNGGNTNGTNGGACNTTYGAYMGNGCNGAR 1989
590 ValAspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgG 606
1990 GCNAAVMGNATGATHGAYGCNTNWSNCAYTGYTNCACNACNTAYMGNA 2039
606 uHisLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspH 623
2040 YTAYAAARGARWSNTGGMNGCCNTGYMGNGCCNMGNGNATGCGCNGARG 2089
623 isThrTrpAspHisAla 628
2090 TNWSNTGGGAYCAYGCN 2106
```

seq\_name: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:US-08-941-445A-6

```
seq_documentation_block:
; Sequence 6, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
```







445 GluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheIleGI 461  
 III :IIIIIIIII :::: :IIIIIIIII :IIIIIIIII :IIIIIIIII  
 1623 GAGCGGGCTTTCGGCGTGACAGCAAAATCCCACTGATCGGTTCATCGS 1672  
 461 YargLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIleProG 478  
 IIIIIIIIII: IIIIIIIIII IIIIIIIIII: IIIIIIIIII  
 1673 CAGGCTGGAGAACAGAAGGCCCTGCAGCTCATGGCCGCCCATCCCCG 1722  
 478 LuLeuMetArgLuAspValGlnPheValMetLeuGlySerGlyAspPro 494  
 IIIIIIIIII: IIIIIIIIII IIIIIIIIII: IIIIIIIIII  
 1723 AGCTCATCAGGAGGAGCTCCAGATCGTTCTTGGGTACTGCAAGAAG 1772  
 495 IlePheGluGlyTrpMetArgSerThrGluSerSerTyrLysAspLysPh 511  
 IIIIIIII: :IIIIIIIII :IIIIIIIII :IIIIIIIII :IIIIIIIII  
 1773 AAGTTYCGAAGAGCTGCTCAAGAGCATGAGGAGAAGTATCCGGCGCAAGET 1822  
 511 eArgGlyTrpValGlyPheSerValProValSerHisArgIleThrAlaG 528  
 IIIIIIII: IIIIIIIIII: IIIIIIIIII: IIIIIIIIII: IIIIIIIIII  
 1823 GAGGCGGCTGGTAAGTTCAACGGCGCGCTGTCATCTCATCATGGCCG 1872  
 528 LYSAspIleLeuLeuMetProSerArgPheGluProCysGlyLysAsn 544  
 IIIIIIIIII: :IIIIIIIII :IIIIIIIII :IIIIIIIII :IIIIIIIII  
 1873 GAGCGAGCTGCTCGCGCTCCCAGCGCTCGAGCGCTTGGAAGTCAATC 1922  
 545 GlnLeuTyAlaMetGlnTyrGlyThrValProValValHisGlyThrGI 561  
 IIIIIIIIII: :IIIIIIIII :IIIIIIIII :IIIIIIIII :IIIIIIIII  
 1923 CAGCTGCAGGGGATGAGATACGGAAGCCGCTGTGTGCGCGCTCCACCG 1972  
 561 yGlyLeuArgAspThrValGluThrPheAsnProPheGlyAlaLysGlyG 578  
 IIIIIIII: IIIIIIIIII: IIIIIIIIII: IIIIIIIIII: IIIIIIIIII  
 1973 TGGGCTCGTGACACGGTCATC.....G 1995  
 578 luGluGlyThrGlyTrpAlaPheSerProLeuThrValAsp..... 591  
 IIIIIIII: IIIIIIII: IIIIIIII: IIIIIIII: IIIIIIII  
 1996 AAGCGCAAGACTGGTTCCACATGGCGGCTCTCACGGCTCGACTGCAAGGTG 2045  
 592 .....LysMetLeuTrpAlaLeuArgThrAlaMe 601  
 IIIIIIII: IIIIIIII: IIIIIIII: IIIIIIII: IIIIIIII  
 2046 GTGGAGCCAAGCAGCGCTGAAGAAGGTGGCGGCCACCCTGAAGCGCGCAT 2095  
 601 tSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgG 618  
 IIIIIIII: IIIIIIII: IIIIIIII: IIIIIIII: IIIIIIII  
 2096 CAAGGTCTGCTC.....GGCAGCGCGGCGTACGGAGAGATGGTCAGGAAC 2139  
 618 LywMetThrLysAspHisThrAspHisAlaProSerSer..ThrSerAr 634  
 IIIIIIII: IIIIIIII: IIIIIIII: IIIIIIII: IIIIIIII  
 2140 GCATGNACCAAGACCTCTCTCTGG...RAGGGCGCTGCGAAGAACTGGGAG 2186  
 634 gSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgGly 650  
 IIIIIIII: IIIIIIII: IIIIIIII: IIIIIIII: IIIIIIII  
 2187 ATATGTGCTCTCTGG..GCTTGGCGCTCGCCG...GCAGCGCGCGCGGA 2229



Sequence	Strd Orig	zScore	EScore	Len	Documentation
/cgn2_6/ptodata/1/pna/US0937_COMB	seq:US-09-345-214-5	+2826.50	2927.34	1.0e-153	2491
/cgn2_6/ptodata/1/pna/US097B_COMB	seq:US-09-743-980-5	+2826.50	2927.34	1.0e-153	2491
/cgn2_6/ptodata/1/pna/US0609_COMB	seq:US-60-094-436-9	+2826.50	2927.34	1.0e-153	2491
/cgn2_6/ptodata/1/pna/US096C_COMB	seq:US-60-094-436-9	+2826.50	2927.34	1.0e-153	2491
/cgn2_6/ptodata/1/pna/US096E_COMB	seq:US-09-654-617-385863	+2821.50	2921.77	2.1e-153	
/cgn2_6/ptodata/1/pna/US096E_COMB	seq:US-09-684-016-385863	+2821.50	2921.77	2.1e-153	
/cgn2_6/ptodata/1/pna/US093_COMB	seq:US-09-345-214-12	+2806.50	2921.77	2.1e-153	
/cgn2_6/ptodata/1/pna/US097B_COMB	seq:US-09-743-980-12	+2806.50	2908.38	1.2e-152	2008
/cgn2_6/ptodata/1/pna/US0609_COMB	seq:US-60-094-436-12	+2806.50	2908.38	1.2e-152	2008
/cgn2_6/ptodata/1/pna/US0934_COMB	seq:US-60-094-436-12	+2806.50	2908.38	1.2e-152	2008
/cgn2_6/ptodata/1/pna/US0936B_COMB	seq:US-09-402-254-52	+2713.50	2812.99	2.4e-147	1749
/cgn2_6/ptodata/1/pna/US0936B_COMB	seq:US-09-673-406-12	+2713.50	2812.97	2.4e-147	1755
/cgn2_6/ptodata/1/pna/US0930_COMB	seq:US-09-627-564-1	+2655.50	2748.28	9.8e-144	2992
/cgn2_6/ptodata/1/pna/US096C_COMB	seq:US-09-654-617-451753	+2578.00	2670.33	2.1e-139	
/cgn2_6/ptodata/1/pna/US096E_COMB	seq:US-09-684-016-451753	+2578.00	2670.33	2.1e-139	
/cgn2_6/ptodata/1/pna/US0936B_COMB	seq:US-09-625-406-20	+2546.50	2640.25	1.0e-137	1620
/cgn2_6/ptodata/1/pna/US0933_COMB	seq:US-09-345-214-6	+2495.00	2587.28	9.1e-135	1528
/cgn2_6/ptodata/1/pna/US097B_COMB	seq:US-09-743-980-6	+2495.00	2587.28	9.1e-135	1528
/cgn2_6/ptodata/1/pna/US0609_COMB	seq:US-60-094-436-10	+2495.00	2587.28	9.1e-135	1528
/cgn2_6/ptodata/1/pna/US06031_COMB	seq:US-60-312-544-3744	+2424.50	2512.47	1.3e-130	1810
/cgn2_6/ptodata/1/pna/US0931_COMB	seq:US-09-345-214-11	+2036.00	2111.40	2.9e-108	1415
/cgn2_6/ptodata/1/pna/US097B_COMB	seq:US-09-743-980-11	+2036.00	2111.40	2.9e-108	1415
/cgn2_6/ptodata/1/pna/US0609_COMB	seq:US-60-094-436-11	+2036.00	2111.40	2.9e-108	1415
/cgn2_6/ptodata/1/pna/US096A_COMB	seq:US-09-606-304-9	+2001.00	2070.78	5.3e-106	2360
/cgn2_6/ptodata/1/pna/US096A_COMB	seq:US-09-606-304-3	+1731.50	1793.46	1.5e-90	1758
/cgn2_6/ptodata/1/pna/US096B_COMB	seq:US-09-620-392-7351	-1523.00	1558.46	1.8e-77	1810
/cgn2_6/ptodata/1/pna/US097A_COMB	seq:US-09-702-134-7216	-1523.00	1558.46	1.8e-77	1810
/cgn2_6/ptodata/1/pna/US095B_COMB	seq:US-09-534-859-7735	+1517.50	1538.90	2.3e-76	8419
/cgn2_6/ptodata/1/pna/US098_COMB	seq:US-09-803-736-735	+1517.50	1538.90	2.3e-76	8419
/cgn2_6/ptodata/1/pna/US096C_COMB	seq:US-09-654-617-258645	+1404.50	1455.51	9.9e-72	
/cgn2_6/ptodata/1/pna/US096E_COMB	seq:US-09-684-016-258645	+1404.50	1455.51	9.9e-72	
/cgn2_6/ptodata/1/pna/US06031_COMB	seq:US-60-312-544-395	+1397.50	1449.23	2.2e-71	1300
/cgn2_6/ptodata/1/pna/US094_COMB	seq:US-09-402-254-48	+1253.50	1394.52	9.4e-63	2423
/cgn2_6/ptodata/1/pna/US096B_COMB	seq:US-09-625-406-10	+1246.00	1287.95	2.1e-62	2097
/cgn2_6/ptodata/1/pna/US094_COMB	seq:US-09-402-254-50	+1239.50	1281.55	4.9e-62	2010
/cgn2_6/ptodata/1/pna/US097B_COMB	seq:US-09-743-980-20	+1239.50	1280.62	5.5e-62	2248
/cgn2_6/ptodata/1/pna/US096B_COMB	seq:US-09-654-617-269739	+1223.00	1269.99	2.1e-61	2248
/cgn2_6/ptodata/1/pna/US096E_COMB	seq:US-09-684-016-269739	+1223.00	1269.99	2.1e-61	
/cgn2_6/ptodata/1/pna/US096A_COMB	seq:US-09-606-304-7	+1215.00	1253.37	1.8e-60	2793
/cgn2_6/ptodata/1/pna/US0934_COMB	seq:US-60-094-436-15	+1212.00	1252.97	1.9e-60	2019
/cgn2_6/ptodata/1/pna/US0971_COMB	seq:US-09-743-980-15	+1212.00	1252.97	1.9e-60	2019



524 GAGCTAAAGTAACACAAAGCATTTGCTTTGTAAACCGCGAAGCTTCTCCT 573  
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167  
574 TATGCAAGCTCTGGGGTCTAGGAGATGTTTGTGGTTCAATTCGCCAGTTCG 623  
167 aLeuAlaAlaArgGlyHlsArgValMetValValMetProArgTyrLeuA 184  
624 TCTTGCTCTCGTGGTCACTGGTGGTGTATGATGCCAGATATTAA 673  
184 sngLysSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHls 200  
674 ATGGTACCTCCGATGAAGAATTATGCAAAATGCAATTTACACAGAAAAACAC 723  
201 lLysIleProCysPheGlyGlySerHlsGluValThrPhePheHlsG 217  
724 ATTCGATTTCCATGCTTTGGGGGTGAACATGAAGTTACCTTCTCCATGA 773  
217 uTyrArgAspAsnValAspTTPValPheValAspHisProSerTyrHlsA 234  
774 GTATAGACATTCAGTTGACTGGTGTGTTGATCATCCCTCATATCACA 823  
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250  
824 GACCTGGAAATTTATATGAGATAAGTTTGGTGTCTTTGGTGATAATCAG 873  
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267  
874 TTCAGATACACACTCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923  
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsna 284  
924 TGAATTTGGGAGGATATTTATGAGACAGAAATGTCATGTTGTGTCAATG 973  
284 sPrPheHlsSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300  
974 ATTGGCATGCCAGTCTAGTGGCAGTCTCTTCTTGTGCAAAATATAGACCA 1023  
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHlsAsnLeuAl 317  
1024 TATGGTGTATTAAGACTCCCGCAGCATCTTGTATATACATAATTTAGC 1073  
317 aHlsGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProP 334  
1074 ACATCAGGGTGTAGAGCCTGCAAGCACATATCTGACCTTGGGTGGCCAC 1123  
334 roGluTrpTyrGlyAlaLeuGluTTPValPheProGluTrpAlaArgArg 350  
1124 CTGAATGGGTATGAGCTCTGGAGTGGGTATTCCTGAAATGGCGGAGGAG 1173  
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367  
1174 CATGCCCTTGACAAGGGTGGAGCAGTTAATTTTGAAGGTGCAGTTCT 1223  
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384  
1224 GACAGCAGATCGAATCGTACTGTCTAGTAAGGGTTATTCTCTGGAGGTCA 1273  
384 hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400  
1274 CAACTGCTGAGGTGGCAGGGCTCAATGAGCTCTTAAGCTCCAGAAAG 1323  
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417  
1324 AGTGATTAACCGAATTTGTAATGGAATTTGACATTAATGATTTGGAACCC 1373  
417 oThrThrAspLysCysLeuProHlsHlsTyrSerValAspAspLeuSerC 434  
1374 TGGCAGACAGAAATGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1423  
434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450  
1424 GAAAGGCCAATGTAAGGTGCATTCGACAGAGAGCTGGGTTTACCTATA 1473

451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467  
1474 AGCCTGATGTTCTCTGATTTGGTATTTATTTGAAGGTTCGATTATCAGAA 1523  
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484  
1524 AGCATTGATCTTCAACTTATCATTACCATCATCTCATCGCGGAAGATG 1573  
484 aGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTyrMet 500  
1574 TTCAATTTGTCATGCTGGATCTGGTGACCCAGAGCTTGAAGATTGGATG 1623  
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTyrValGlyPh 517  
1624 AGATCTACAGATCGATCTTCAAGGATAAATTTCTGTTGGTGGTGGATTT 1673  
517 eSerValProValSerHlsArgIleThrAlaGlyCysAspIleLeuLeu 534  
1674 TAGTGTTCAGTTTCCACCGAATAACTGCCGGCTGCCGATATATTGTTAA 1723  
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550  
1724 TGCCATCCAGATTCGAACCTTGGTCTCAATCAGCTATATGCTATGCAG 1773  
551 TyrGlyThrValProValValHlsGlyThrGlyGlyLeuArgaspThrVa 567  
1774 TATGCCAGATTCCTGTTTCCATGCAACTGGGGGCTTTAGAGATACCGT 1823  
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584  
1824 GGAGAACTTCAACCTTTCCGTTGAGATGAGAGCAGGTTACAGGTGG 1873  
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600  
1874 CATTCGCACCCCTAACACACAGAAACATGTTGTGGACATTCGGAACCTGA 1923  
601 MetSerThrPheArgGluHlsLysProSerTrpGluGlyLeuMetLysAr 617  
1924 ATATCTACATACAGGAACACAACTCTCTCTGGGAAGGGCTTAATGAAGCG 1973  
617 gGlyMetThrLysAspHlsThrTrpAspHlsAlaProSerSerThrSer 633  
1974 AGGATGTCAAAAGACTTCAGCTGGGACCATGCCGCTGAACAAATACGAAC 2023  
634 ArgSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG 650  
2024 AAATCTCCAGTGGGCTTCATCGATCGACCTATGTCATGTAATAAAA... 2071  
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667  
2072 .....AGGACCAAAAGTGGTTCCT.....TCAAGAT 2099  
667 erSerPheArgGlyPro.....GluGlyTyrProCys 677  
2100 CATCAGTTTCATCATCTTATAGTAACTAAATGATGAAAGAAAAACCCCTGT 2149  
678 ThrLeuArgCysProAlaThrValGluSerGlnCysAlaCysLeuLeuTr 694  
2150 ACATTTACAT..... 2158  
694 pPheAlaGlySerArgThrTyrAspGlyCysAlaAlaAlaValThrA 711  
2158 ..... 2158  
711 laSerGlyGlyArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAla 727  
2159 .....GGAAGGCAGACCGCTATTGGCTCCATTTGCTCCAACGCTGTC 2200  
728 AlaGlyTrpLeuThrAlaLysHls 735  
2201 TTTGGCTGGCTTGGCTTCGATGCAC 2224











201	lIeLysIleProCysPheGlySerHisGluValThrPhePheHisG1	2117
724	ATTCGGATTCCATGCTTTGGCGGTGAACATGAAGTTACCTTCTTCATCA	773
217	uTyArgAspAsnValAspTrpValPheValAspHisProSerTyHisA	234
774	GTATAGAGATTCCAGTTGACTGGGTGTTTGTGATCATCCCTCATATCA	823
234	rpProGlySerLeuTyTyGlyAspAsnPheGlyAlaPheGlyAspAsnGln	250
824	GACCTCGGAATTTATATGCGAGATAGTTTGGTGTCTTTGGTGATAATC	873
251	PheArgTyThrLeuLeuCysTyAlaAlaCysGluAlaProLeuIleLe	267
874	TTACAGATACACATCTTTGTGTATGCTGCATGTGAGGCTCTTTGATCC	923
267	uGluLeuGlyGlyTyIleTyGlyGlnAsnCysMetPheValValAsnA	284
924	TGAATTTGGGAGGATATATTTATGACAGAAATTCATGATGTTGTTCT	973
284	spTrpHisAlaSerLeuValProValLeuLeuAlaLysTyArgPro	300
974	ATTGCGATGCCAGTCTAGTGGCAGTCCCTCTGTGTCGAAATATATAG	1023
301	TyrGlyValTyArgAspSerArgSerThrLeuValIleHisAsnLeuAl	317
1024	TATGTGTTTATAAAGACTCCCGCAGCATCTTCTGATACATAATTTAG	1073
317	aHisGlnGlyValGluProAlaSerThrTyTyProAspLeuGlyLeuPro	334
1074	ACATCAGGGTGTAGAGCCTGCAAGCACATATCTTGACCTTGGTGTGC	1123
334	roGluTrpTyGlyAlaLeuGluTrpValPheProGluTrpAlaAaArg	350
1124	CTGAATGATATGAGCTCTGGAGTGGGTATTCCTGAATGGCGAGGAG	1173
351	HisAlaLeuAspLysGlyGluAlaValAspPheLeuLysGlyAlaVal	367
1174	CATGCCCTTGACAAGGTGAGGCAGTTAAATTTTGAAGAGTGCAGTTG	1223
367	lThrAlaAspArgIleValThrValSerGlnGlyTySerTrpGluVal	384
1224	CACAGCAGATCGAATCGTACTGTCAGTAAGGTTATTCCTGGGGAGGT	1273
384	hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys	400
1274	CAACTGCTGAAGTGGACAGGGCCTCAATGAGCTCTTAAGCTCCAGAA	1323
401	SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr	417
1324	AGTGTATTAAACGGAATGTGAATGGAAATTTGACATTAATGATTGGA	1373
417	oThrThrAspLysCysLeuProHisHisTySerValAspAspLeuSerG	434
1374	TGCCACAGACAAATGTATCCCTCTCATATTCTGTGTGATGACCTCTCG	1423
434	lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal	450
1424	GAAAGGCCAAATGTGAAGGTGCATATGCAGAAGGAGGTGGGTTTAC	1473
451	ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyArgL	467
1474	AGGCCATGATCTTCTCTGATTGGCTTTATTTGGAAGGTTGGATTAT	1523
467	sgLysIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAsp	484
1524	AGGCATTGATCTCATCAACTTATCATACAGATCTCATCGCGGAAGAT	1573
484	alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet	500
1574	TTCAATTTGTCATGCTTGGATCTGGTGACCCAGAGCTTGGAGATTGG	1623
501	ArgSerThrGluSerTyTyLysAspLysPheArgGlyTrpValGlyPh	517

1624 AGATCTACAGATCGATCTTCAAGGATAAATTTCTGGATGGGTGGATT 1673

517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534

1674 TAGTGTTCAGTTTCCACCGAATAACTGCCGGCTGCGATATATTGTTAA 1723

534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyAlaMetGln 550

1724 TGCCATCCAGATCGAACCTTGCTGTCTCAATCAGCTATATGCTATCGAG 1773

551 TyrGlyThrValProValHisGlyThrGlyGlyLeuArgAspThrVa 567

1774 TATGGCACAGTTTCCTGTTGTCATCAACTGGGGCCCTTAGAGATACCGT 1823

567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584

1824 GGAGAACTTCAACCCCTTTCGGTGAAGATGGAGACAGGGTACAGGGTGG 1873

584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600

1874 CATTCGCACCCCTAACCCACAGAAAAACATGTTGTGGACATTCCGAACGTCA 1923

601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617

1924 ATATCTACATACAGGGAACACAAGTCTCTCTGGGAAGGGCTAATGAAGCG 1973

617 gGlyMetThrLysAspHisThrTrpAspHisAlapro\_SerSerThrSer 633

1974 AGGCATGTCAAAAGACTTCACGTGGGACCATGCCGTGAACAATACGAAC 2023

634 ArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgGl 650

2024 AAATCTTCAGTGGGCCCTTCATCGATCGACCTATGTCATGTATGAATAAA 2071

650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667

2072 .....AGGACCAAGGGTGGTTCCT.....TGAAGAT 2099

667 eSerPheArgGlyPro.....GluGlyTyProCys 677

2100 CATTCAGTTTCATCTCTATAGCTAAGCTAATATGATGAAGAAAAACCCGTGT 2149

678 ThrLeuArgCysProAlaThrValGluSerGlnCysAlaCysLeuLeuTr 694

2150 ACATTACAT..... 2158

694 pPheAlaGlySerArgThrTyAspGlyCysAlaAlaAlaValThrA 711

2158 ..... 2158

711 laSerGlyGlyArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAla 727

2159 .....GGAGCGACAGCCGCTATTGGCTCCATTCCTCCACGTCGTGC 2200

728 AlaGlyTrpLeuThrAlaLysHis 735

2201 TTTGGCTGGCTTGCCTCGATGCAC 2224

seq\_name: /cgn2\_6/ptodata/1/pna/US096C\_COMB.seq:US-09-654-617-385863

```

seq_documentation_block:
  : Sequence 385863, Application US/09654617
  : GENERAL INFORMATION:
  :
  : APPLICANT: Kovalic, David K.
  :
  : APPLICANT: Liu, Jingdong
  :
  : TITLE OF INVENTION: Annotated Plant Genes
  :
  : FILE REFERENCE: 38-21(15097)D
  :
  : CURRENT APPLICATION NUMBER: US/09/654,617
  :
  : CURRENT FILING DATE: 2000-09-05
  :
  : NUMBER OF SEQ ID NOS: 463173
  :
  : SEQ ID NO 385863
  :
  : LENGTH: 2606
  :

```



```
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-654-617-385863
```

```
alignment_scores:
  Quality: 2821.50      Length: 726
  Ratio: 4.522          Gaps: 18
  Percent Similarity: 85.950  Percent Identity: 76.309
```

```
alignment_block:
US-09-674-824-2 x us-09-654-617-385863 ..
```

```
Align seg 1/1 to: US-09-654-617-385863 from: 1 to: 2606
```

```
2 AlaAlaThrGlyValGlyAlaGlyCysLeu...AlaProSerValArgLe 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 CCGGGGATGGGGATCGGGCGCGCTGCTGGTGGCGCGCAGCGTGAGG... 198
17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199 .....CGGGGAGG.....AGGT 211
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluSerArg 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
212 TCGCGCTCCAGCGGTGGGAGG...CGGTGCGTGGCGGAGCTGAGCAGG 258
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaPro..... 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
259 GACGGTGGGTGGGCG.....CAGCGCGCTGGGACCGCGGCC 296
66 .....ProLeuValProGlyPheLeuAlaPro.....Prop 76
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 GCTGGTGAAGCAGCGCGTCTCGCGACCTCTCTGTCGCGCAGCTCGAGCG 346
76 roProAlaProAlaGlnSer.....ProAlaProThrGlnProProLeu 90
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
347 CACCGCGCGCCACGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGTGG 396
91 ProAspAlaGlyValGlyGluLeuAlaProAspLeuLeuGluGlyI 107
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
397 CCGGACTCCGCGGTGGGAGATCAGCGCCGAT.....CTAGAGGTCT 440
107 eaGluAspSerIleAspSerIleValAlaAlaSerGluGlnAsps 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
441 CACAGAAGATTCATCGACAAACAAATTTTGTGGCTAGTGAGCAGGAGT 490
124 erGluIleMetAspAlaAsnGluGlnProGlnAlaLysValThrArgSer 140
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
491 CTGAGATCATGGATGTGAAGGAGCAAGCTCAAGCTAAAGTACACGCCAGC 540
141 IleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLe 157
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
541 GTTGCTTTGTAAACCGGTCAACCTCTCTTATGCAAAAGTCAGGTGGACT 590
157 uGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHisA 174
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
591 AGGAGATGTTTGTGGTTCACTGCCAAATTTGCTCTCTCTGCTGCTATC 640
174 rgValMetValValMetProArgTyrLeuAsnGlySerSerAspLysAsn 190
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
641 GTGTGATGTTGTATGCGGAGATACATGAACGGCGCTTGACAAAAT 690
191 TyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheGl 207
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
691 TTTGCAACCGCATTTTACACTCAGAAGCACATTAAGATTCCATGCTTTGG 740
207 yGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValAspT 224
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
741 CGGAGAACATGAAGTACTATTTTTCACGAGTATAGGATTCCTGTTGAT 790
224 rpValPheValAspHisProSerTyrHisArgProGlySerLeuTyrGly 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
791 GGGTGTTTGTGATCATCCCTCATATCATAGACCTCGAAATTTGTATGCA 840
241 AspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCy 257
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
841 GATAAATTTGGTGCITTTGGCGATAATCAGTTCAGATACACACTCCTGTG 890
257 sTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleT 274
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
891 CTATGCGCGGTGTAAGCCCAATTAATCTTGAACCTGGGAGGATATATCT 940
274 yrGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuVal 290
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
941 ATGGACAGAAATGCATGTTTGTGTGAATGATGGCATGCCACTCTGTGG 990
291 ProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSe 307
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
991 CCACTCTCTCTCTGCTGCAAAATATACCCATATGTTGTTTACAGGGATGC 1040
307 rArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProA 324
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1041 CCGCAGTGTCTTGTCTATACATAATCTAGCACATCAGGGTGTGGAGCCTG 1090
324 laSerThrTyrProAspLeuGlyLeuProGluTrpTyrGlyAlaLeu 340
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1091 CCACTACATATCTGACTGCGGATTCGCCACCTGAATGGTATGCGACATTA 1140
341 GluTrpValPheProGluTrpAlaArgHisAlaLeuAspLysGlyGI 357
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1141 GAATGGGTGTTTCCAGAGTGGCAAGCGGCATGCCCTTGACAAGGGTGA 1190
357 uAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValT 374
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1191 GGCAGTCAATTTTAAAGCGCGAGTTGTGACAGCAGATCGAATTTGTGA 1240
374 hrValSerGlnGlyTyrSerTrpGluValThrThrAlaGluGlyGln 390
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1241 CTGTCACCCAGGGGTATTATGGGAGGTCACAACTGCTGAAGGTGGCAA 1290
391 GlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleVa 407
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1291 GGCCTCAATGAGCTCTTAAAGCTCCCGAAGAGTGTATTGAATGGAATGT 1340
407 laSnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuP 424
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1341 AAATGGAAATGACATTAATGATGGAAACCCATCCACAGACAAAGTTCTCC 1390
424 roHisHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysAla 440
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1391 CTTATCATTTATCTGTTGATGACCTGTCCGAAAGGCCAAAGTGTAAAGCT 1440
441 GluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuII 457
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1441 GAATTCGAGAAGGAGCTGGGTTTACCTATGAAGCCCGCATGTGCCCTCTGAT 1490
457 eGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysM 474
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1491 TGGCTTTATTCGAAGATTTGACTATCAAAAAGGCATTCATCTAATTAAC 1540
474 etAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuGly 490
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1541 TTGCCATTCACAGATCTCATCGGGACAATATTCAATTCGTCATGCTTGGGA 1590
491 SerGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSerty 507
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1591 TCTGGTGACCCAGGTTTTTGAAGATGGATGAGATCCACAGANTCAGGGTA 1640
507 rLysAspLysPheArgGlyTrpValGlyPheSerValProValSerHisA 524
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1641 CAGGGATAAATTTCTGCGATGGGTGGGATTAGTGTCCAGTTTCCACCACC 1690
524 rgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluPro 540
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1691 GAATAACTGCAGGTTGCCATATATATTGTTGATGCCATCCAGATTCCAACT 1740
```



```

541 CysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValVa 557
|||||
1741 TGTGGCTCAATCAGCTATATGCTATGCAATATGGTACAGTGCCTGTTGT 1790
557 lHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheG 574
|||||
1791 TCATGAACTGGAGGCTCAGAGATACAGTGAGAAATTTTAAACCGTTTG 1840
574 lAlaLysGlyGluGluGlyThrGlyTyrAlaPheSerProLeuThrVal 590
::: |||||
1841 CTGAGAAAGAGAGAGGCTACAGGTGGGCATTCTCGCCACTAACCATTT 1890
591 Asp.LysMetLeuTyrAlaLeuArgThrAlaMetSerThrPheArgGluH 607
::: |||||
1891 GAAAAAATGCTGTGGGCATTCTCGGCATTTTCGACATACAGGAAC 1940
607 lAspProSerThrGluGlyLeuMetLysArgGlyMetThrLysAspHis 623
|||||
1941 ACAAGTCTCTTGGAGGCTCTAATGAAGGAGGATGTCAGGCGACTTT 1990
624 ThrTrpAspHisAla.ProSerSerThrSerArgSerSerGlyProS 640
|||||
1991 ACATGGGACCATGCGCCTCACAGTATGAACAGATCTTCGAATGGGCTT 2040
640 erTrpThrAsnProThrSerCysArg.ArgGlyLeuGlyArgSerLys... 655
|||||
2041 CATGATCAACCATATGTCATGTAATGATTTGAAGGAGCAGCAATT 2090
656 .CysGluSerPro.....SerAlaLeuLysThrS 665
|||||
2091 TCTCCAGAGCCCTCAATCTCTCTCTCTTCATGAGCGGAATGAAACTT 2140
665 erSerSerSerPhe.ArgGly.....Pr 672
2141 TGTACACTACATGGAAGGGAACCACTTATGCAAGTTGCAACAGATCAC 2190
672 oGluGlyTyrProCysThrLeuArgCysProAlaThrValGluSerGlnC 689
|||||
2191 TCAGGTATACCTTGTAGG.....CCTGCTACTTGGCCA..... 2224
689 ysAlaCysLeuLeuTrpPhe 695
::: |||||
2225 .....ATATGGTTC 2233

```

seq\_name: /cgn2\_6/ptodata/1/pna/us096E\_COMB.seq:us-09-684-016-385863

```

seq_documentation_block:
; Sequence 385863, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 385863
; LENGTH: 2606
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-684-016-385863

```

alignment\_scores:  
Quality: 2821.50 Length: 726  
Ratio: 4.522 Gaps: 18  
Percent Similarity: 85.950 Percent Identity: 76.309

alignment\_block:

US-09-674-824-2 x US-09-684-016-385863 ..

```

Align seg 1/1 to: US-09-684-016-385863 from: 1 to: 2606
2 AlaAlaThrGlyValGlyAlaGlyCysLeu...AlaProSerValArgLe 17
|||||
151 GCGGGATGGGATCGGGCGCGCTGCTGCTGGCGCGCAGGTGAGG... 198
17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
|||||
199 .....CCGGGAGG.....AGGT 211
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
|||||
212 TCGGGCTCCACGCGGTGCGGAGG...CGGTGCTGGCGGAGCTGACGAG 258
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaPro..... 65
::: |||||
259 GACGCTGGGTGGCG.....CACGGCCGCTGGCACCGCGGCC 296
66 .....ProLeuValProGlyPheLeuAlaPro.....Prop 76
|||||
297 GCTGGTGAAGCAGCGCTGCTGCGGACCTTCTCGTCCGCGACGTGCGACG 346
76 roProAlaProAlaGlnSer.....ProAlaProThrGlnProProLeu 90
|||||
347 CACCGCGCCACGACGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 396
91 ProAspAlaGlyValGlyGluLeuAlaProAspLeuLeuLeuGluGlyII 107
|||||
397 CCGGACTCCGCGTGGGGAGATCGAGCCCGAT.....CTAGAAGGTCT 440
107 eAlaGluAspSerIleAspSerIleValAlaAlaSerGluGlnAspS 124
|||||
441 CACAGAGATTCATCGACACAAACAAATTTTGTGCTAGTGACGAGAGT 490
124 erGluIleMetAspAlaAsnGluGlnProGlnAlaLysValThrArgSer 140
|||||
491 CTGAGATCATGGATGTGAGGAGCAGCAAGCTTCTCTATGCAAGTCAAGT 540
141 lIleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLe 157
|||||
541 GTTGTCTTTGTAACCGGTGAAGCTTCTCTTATGCAAGTCAAGTCAAGT 590
157 uGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHis 174
|||||
591 AGGAGATGTTGGTTCACATGCCAATGCTCTTCTCTCTGCTGCTGTCATC 640
174 rgValMetValMetProArgTyrLeuAsnGlySerSerAspLysAsn 190
|||||
641 GTCTGATGCTTGTATGCGGAGATACATGACGGGCGCTTGACAAAT 690
191 TyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheG 207
|||||
691 TTTGCAACGCAATTTTACATCAGAGCAGCATTAAAGATTCCATGCTTGG 740
207 yGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValAspT 224
|||||
741 CGGAGAACATGAAGTACTTTTTTCACGAGTATAGGGAATCTGTGATT 790
224 rpValPheValAspHisProSerTyrHisArgProGlySerLeuTyrGly 240
|||||
791 GGGTGTGTTGATCATCTCATATCATAGACCTGGAAATTTGATGGA 840
241 AspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCy 257
|||||
841 GATAATTTTGGTCTTTTGGCGATTAATCAGTTCAGATACACACTCCTGTG 890
257 sTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyTyrIleT 274
|||||
891 CTATCGCGGTGTGAAGCCCAATTAATTCCTTGAACCTGGGAGGATATATCT 940
274 yGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuVal 290
|||||

```



941 ATGGACAAATGATGTTTCTTGTGAATGATGGCATGCCAGTCTTGTG 990  
291 ProValLeuLeuAlaLysTyrArgProTyrGlyValTyrArgAspSe 307  
991 CCAGTCCCTTCTGTGCAAAATATAGACCATATGTTGTACAGGGATGC 1040  
307 rArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProA 324  
1041 CGCGAGTGTCTTGTTCATACATAATCTAGCACATCAGGGTGTGGAGCCTG 1090  
324 laSerThyTyrProAspLeuGlyLeuProProGluTyrTrpTyrGlyAlaLeu 340  
1091 CCAGTACATATCTGACCTGGGATGGCCACCTGAATGGTATGGAGCATTA 1140  
341 GluTrpValPheProGluTrpAlaArgArgHisAlaLeuAspLysGlyG 357  
1141 GAATGGGTGTTTCCAGAGTGGGCAAGGCGGATGCCCTTGACAGGGTGA 1190  
357 uAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValT 374  
1191 GCGAGTCAATTTTAAAGGCGCAGTGTGACAGCACATCGAATTTGTGA 1240  
374 hrValSerGlnGlyTyrSerTrpGluValThrAlaGluGlyGln 390  
1241 CTGTGAGCCAGGGGTATTTCATGGGAGGTACAACTGCTGAAGGTGGGCAA 1290  
391 GlyLeuAsnGluLeuSerSerArgLysSerValLeuAsnGlyIleVa 407  
1291 GGCCTCAATGAGCTCTTAAGCTCCCGAAGAGATGTATTGAATGGAAATGT 1340  
407 lAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuP 424  
1341 AAATGGAATGACATTAATGATTGGAACCCATCCACAGACAAGTTTCTCC 1390  
424 roHisHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysAla 440  
1391 CTTATCATTTATTCGTGTGATGACCTGTCCGGAAGGCCAAGTGTAAAGCT 1440  
441 GluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuI 457  
1441 GAATTCGACAGAGAGTGGTGTACCTATTAAGCCCGATGTCCTCTGAT 1490  
457 eGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysM 474  
1491 TGGCTTTATTTGGAAGATTGGACTATCAAAAGGCATTTGATTAATAAAC 1540  
474 etAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuGly 490  
1541 TTGCCATTTCCAGATCTCATCGGGGACAAATATCAATTCGTCTGCTGGA 1590  
491 SerGlyAspProPheGluGlyTyrMetArgSerThrGluSerSerTy 507  
1591 TCTGGTCACCCAGTTTGAAGGATGGATGAGATCCACAGAAATCAGGGTA 1640  
507 rLysAspLysPheArgGlyTrpValGlyPheSerValProValSerHisA 524  
1641 CAGGGATAAATTCGTGGATGGGTGGATTAGTTTACGATTTCACCTCCACC 1690  
524 rgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluPro 540  
1691 GAATAACTGCGAGTGGCATATATTGTTGATGCCATCCAGATTCGAACCT 1740  
541 CysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValVa 557  
1741 TGTGGCTCAATCAGCTATATGCTATGCAATATGTTACAGTGCCTGTGTG 1790  
557 lHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheG 574  
1791 TCATGGAACTGGAGCGCTCAGAGATACAGTGGAGAATTTTAACCGTTTG 1840  
574 lyAlaLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrVal 590  
1841 CTGAGAAAGGAGACAGGGTACAGGGTGGGCATTTCTGCCACTAACCAT 1890

591 Asp.LysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluH 607  
1891 GAAAAAATGCTGTGGGCATTCGGATGCAATTTTCACATACAGGAAC 1940  
607 isLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHis 623  
1941 ACAAGTCCCTCTTGGGAGGTCTAATGAAGCAGGCGCATGTCAAGTGGCCTT 1990  
624 ThrTrpAspHisAla.ProSerSerThrSerArgSerSerSerGlyProS 640  
1991 ACATGGGACCATGCCCGCTCACAGTATGAACAGATCTTCAATGGGCCTT 2040  
640 erTrpThrAsnProThrSerCysArg.ArgGlyLeuGlyArgSerLys.. 655  
2041 CATGGATCAACCATATCTCATGTAAATGGATTGAAAGGAGCAGCGAAAT 2090  
656 CysGluSerPro.....SerAlaLeuLysThrS 665  
2091 TCTCCGAGGACCTCAATCTCTCTCTCTTCATGAGCGGAATGAAACTT 2140  
665 erSerSerSerPhe.ArgGly.....Pr 672  
2141 TGTACACTACATGAAAGGCAACCAAGTTATGCAAAAGTTGCAACGATCAC 2190  
672 oGluGlyTyrProCysThrLeuArgCysProAlaThrValGluSerGlnC 689  
2191 TCAAGGTACCCCTTGTAGG.....CCTGCTACTTGGCCA..... 2224  
689 ysAlaCysLeuLeuTrpPhe 695  
2225 .....ATATGGTTC 2233

seq\_name: /cgn2\_6/ptodata/1/pna/US093\_COMB.seq:us-09-345-214-12

seq\_documentation\_block:

; Sequence 12, Application US/09345214  
; GENERAL INFORMATION:  
; APPLICANT: Lightner, Jonathan E.  
; APPLICANT: Broglie, Karen E.  
; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE  
; TITLE OF INVENTION: EXPRESSION TO PRODUCE STARCHES IN GRAIN CROPS  
; FILE REFERENCE: BB-1147  
; CURRENT APPLICATION NUMBER: US/09/345,214  
; CURRENT FILING DATE: 1999-06-30  
; EARLIER APPLICATION NUMBER: 060/094,436  
; EARLIER FILING DATE: 1998-07-28  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 12  
; LENGTH: 2008  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-345-214-12

alignment\_scores:  
Quality: 2806.50 Length: 660  
Ratio: 4.717 Gaps: 6  
Percent Similarity: 90.152 Percent Identity: 80.909

alignment\_block:

US-09-674-824-2 x US-09-345-214-12 ..

Align seg 1/1 to: US-09-345-214-12 from: 1 to: 2008

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20  
:.....:|||||

14 TCGGCGGTGGGCGCGCGTGCCTCTCTC..... 40

20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34  
:.....:|||||

41 .CTCGCGCGGCGCGCTGCGCGCGCGCGCTCGCGCGCGCGCGCGCGC 89  
:.....:|||||



[illegible]

334	roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg	350
981	CTGAATGGTATGGAGCTCTGGAGTGGGTATTCCCTGAATGGCGAGGAGG	1030
351	HisAlaLeuAspLysGlyGluAlaValAlaAsnPheLeuLysGlyAlaValVa	367
1031	CATGCCCTTCACAAGGGTGAGGCAGTAAATTTTGTAAAGGTGCAGTTGT	1080
367	lThrAlaAspArgIleValThrValSerClnGlyTyrSerTrpGluValT	384
1081	GACAGCAGATCGAATCGTCACTGTCTAGTAAGGGTTATTCTGGTGGAGGTCA	1130
384	hrThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys	400
1131	CAACTGCTCAAGGTGACAGGGCTCAATGAGCTCTTAAGCTCCAGAAG	1180
401	SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr	417
1181	AGTGTAATAACCGGAATTCGAATGGAATTCACATTAATGATGGAAACC	1230
417	oThrAspLysCysLeuProHisIleTyrSerValAspAspLeuSerG	434
1231	TGCCACAGACAAATGTATCCCTGTCATTAATTCCTGTTGATGACCTCTCTG	1280
434	lYlysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal	450
1281	GAAGGCCAATGTAAAGGTGCATTGCCAGAAGAGCTGGGTGTACCTATA	1330
451	ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnL	467
1331	AGGCTGTGATGTCCCTCTGATTGGCTTTATTGGGAAGGTGGATTATCAGAA	1380
467	sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV	484
1381	AGGCATTCGATCTCATTCAACTATTATCATACAGATTCATGCGGGAAGATG	1430
484	alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet	500
1431	TTCAATTTGTATGCTTGGATCTGGTCACCCAGAGCTTGAAGATTGGATG	1480
501	ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh	517
1481	AGATCTACAGACTCGATCTCTCAGGATAAATTTCTGGTGATGGGTGGATT	1530
517	eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM	534
1531	TAGTGTTCAGTTTCCACCAGAAATACTGCCGCTCCGATATATTGTAA	1580
534	eProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln	550
1581	TGCCATCCAGATTGGAACCTTGTGGTCTCAATCAGCATATATGCTATG	1630
551	TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa	567
1631	TATGGCAGATTCTGTTGCCATGCAACTGGGGCCCTTAGAGATACCGT	1680
567	lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA	584
1681	GGAGAACTTCAACCTTTCCGTGAGATGGAGACAGGGTACAGGGTGGG	1730
584	laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla	600
1731	CATTGCCACCCCTTAACCACAGAAAACATGTTGTGGACATTCGGAAC	1780
601	MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr	617
1781	ATATCTACATACAGGGAACACAAAGTCTCTCTGGGAAGGCCTAATGAAGCG	1830
617	gGlyMetThrLysAspHisIleThrTrpAspHisAlaPro_SerSerThrSer	633
1831	AGGCATGTCAAAGACTTCCACGTGGGACCATGCCCTGTAACAAATACCAAC	1880
634	ArgSerSerSerGlyProSerThrTrpThrAsnProThrSerCysArgArgG	650



Fri Mar 29 09:26:44 2002

```

1881  AAATCTTCAGTGGCCCTTCATCGACCCATGTCATGTATAAAAA... 1928
650  yLeuGlyArgSerLysCysGluSerPro 659
1929  .....AGGACCAAGTCGGTTCCT 1949
seq_name: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:us-09-743-980-12

```

```

seq_documentation_block:
; Sequence 12, Application US/09743980
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE
; TITLE OF INVENTION: EXPRESSION TO PRODUCE STARCHES IN GRAIN CROPS
; FILE REFERENCE: BB-1147-A
; CURRENT APPLICATION NUMBER: US/09/743,980
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 060/094,436
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 2008
; TYPE: DNA
; ORGANISM: Zea mays
US-09-743-980-12

```

```

alignment_scores:
Quality: 2806.50      Length: 660
Ratio: 4.717          Gaps: 6
Percent Similarity: 90.152 Percent Identity: 80.909

```

```

alignment_block:
US-09-674-824-2 x US-09-743-980-12 ..
Align seg 1/1 to: US-09-743-980-12 from: 1 to: 2008
4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20
14 TCGGCCGTGGCGCGCGTCCCTC..... 40
20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34
41 .CTCGCGCGCGCGCTGGCGCGCGCGTCCGCGACCGCGCGCGCGC 89
35 ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
90 GCGCGCTCCAGCGGTGCTGCGCGCGCGCGTCCGCGCGCGCGTCCGCG 139
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnGlnGlnGlnGln 67
140 GAGGGCGCGCGCGCGCGCGCGCGTCCCGCGCGCGCGCGCGCGCT 189
67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerPro 84
190 CGTGGCGCGCGCTTCCTCGG...CGCGCGCGCGCGCGCGCGCGCGCG 236
84 laProThrClnProProProProAlaGlyValGlyGluLeuAlaPro 100
237 CATCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 283
101 AspLeuLeuGluGlyLeuAlaGluAspSerLeuAspSerLeuLeuLeu 117
284 ...CTGACCACTGAAGGATTCCTCAAGGTTCCTCAAGTAAACACAGTAGT 330
117 lAlaAlaSerGluGlnAspSerGluLeuMetAspAlaAsnGluGlnPro 134
331 TGTGCAAGTACAGAGATCTGAGATCTGTTGTTGGAAGGACGAGCTC 380
134 lAlaLysValThrArgSerLeuValPheValThrGlyGluAlaPro 150

```

```

381 GAGCTAAAGTAACACAAAGCATGTCTTTGTAAACGGCGAAGCTTCTCCT 430
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
431 TATGCAAGTCTGGGGTCTAGGAGATGTTGTGGTTCATTGCCAGTTGC 480
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
481 TCTTGTCTGCTGCTGCTACCGTGTGATGTTGTAATGCCAGATATTAA 530
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
531 ATGGTACCTCCGATAGAATATATGCAATATGCAATATGCAATATGCA 580
201 lLeLysileProCysPheGlyGlySerHisGluValThrPhePheHisG 217
581 ATTCCGATTCCTGCTTGGCGTGAACATGAATGAATGCTTCTTCCATGA 630
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
631 GTATAGAGATTCAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 680
234 rgProGlySerLeuTyrGlyAspAsnGlyGlyAlaPheGlyAspAsnG 250
681 GACCTGGAAATTTATATGGAGATAAGTTGGTCTTTGGTGTGATAATCA 730
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuLe 267
731 TTCAGATACACACATCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 780
267 uGluLeuGlyGlyTyrTyrTyrGlyGlnAsnCysMetPheValValAsn 284
781 TGAATTTGGAGGATATATTTATGGACAGAAATTCATGTTTGTGTCA 830
284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300
831 ATTGCATGCCAGTCTAGTGCAGTCTCTTCTGCTGCAAAATATAGACCA 880
301 TyrGlyValTyrArgAspSerArgSerThrLeuValLysHisAsnLeu 317
881 TATGGTGTATATAAGTCTCCCGCAGCATCTTCTGTAATACATATTA 930
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuPro 334
931 ACATCAGCGTGTAGAGCTGCAAGCACATATCTGACCTTGGGTGGCAC 980
334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTyrAlaArgArg 350
981 CTGAATGTGATGGAGCTCTGGAGTGGGTATTTCCCTGAATGGCGAGG 1030
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaVal 367
1031 CATGCCCTTGACAGGGTGAGCAGTTAAATTTTGAAGGTGCAGTTGT 1080
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluVal 384
1081 GACAGCAGATCGAATCGTACGTGTCAGTAAAGGTTATTCGTGGGAGG 1130
384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400
1131 CAACTGCTCAAGGTGACAGGCGCTCAATGAGCTCTTAAGCTCCAGAA 1180
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsn 417
1181 AGTGTATTAAACGGAATTTGTAATGGAATTTGACATTAATGATTGGA 1230
417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuS 434
1231 TGCACAGACAAATGATCCCTGCTGCTATTTCTGTTGATGACCTCTCT 1280
434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450
1281 GAAAGCCAAATGTAAGGTGCATTTGCAAGAGGAGCTGGGTTTACCTA 1330

```



```

451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467
1331 AGCGCTGATGTTCTCTGATGGCTTATTGGAAGGTTGGATTATCAGAA 1380
467 sGlyIleAspLeuIleGlyMetAlaIleProGluLeuMetArgGluAspV 484
1381 AGGCATTGATCTCATCACTTATCATACAGATCTCATCGGGAAGATG 1430
484 aGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500
1431 TTCAATTGTCATGCTTGGATCTGGTGACCCAGAGCTTGAAGATTGGATG 1480
501 ArgSerThrGluSerTyrIleAspLysPheArgGlyTrpValGlyPhe 517
1481 AGATCTACAGATCGATCTTCAAGGATAAATTCGTGGATGGTGGATT 1530
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534
1531 TAGTGTTCAGATTTCACCGAATAACTCGCGGCTGCGATATATTGTTAA 1580
534 eTProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550
1581 TGCCATCCAGATTCGAACCTTGTGTCATCACTATATGCTATGCGAG 1630
551 TyrGlyThrValProValHisGlyThrGlyLeuArgAspThrVa 567
1631 TATGGCACAGTTCTCTGTTGCTCATCACTGGGGGCTTAGAGATACCGT 1680
567 lGluThrPheAsnProPheGlyValAlaGlyCysGluGlyThrGlyTrpA 584
1681 GGAGAACTTCAACCTTCCGTGAGAAATGAGGAGGAGGAGGAGGAGG 1730
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600
1731 CATTCGACCCCTAACCCACAGAAACATGTTGTCGACATTGCGAACTGCA 1780
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617
1781 ATATCTACATACAGGGAACAAAGTCCTCCCTGGGAAGGCTTAATGAAGC 1830
617 gGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSer 633
1831 AGGCATGTCAAAGACTTCACGTGGGACCATGCGCTGAACAATACGAAC 1880
634 ArgSerSerGlyProSerTrpThrAsnProThrSerCysArgArgGl 650
1881 AAATCTCCAGTGGGCTTCATCGATCGACCCCTATGTCATGTAATAAAA 1928
650 yLeuGlyArgSerLysCysGluSerPro 659
1929 .....AGGACCAAGTGGTGGTTCCT 1949

```

```
seq_name: /cgn2_5/ptodata/1/pna/us6009_COMB.seq:us-60-094-436-12
```

```

seq_documentation_block:
; Sequence 12, Application US/60094436
; GENERAL INFORMATION:
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Broglie, Karen E.
; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: Modification of Starch Biosynthetic Enzyme Gene
; FILE REFERENCE: BB-1147-p1
; CURRENT APPLICATION NUMBER: US/60/094,436
; CURRENT FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 12
; LENGTH: 2008
; TYPE: DNA
; ORGANISM: Zea mays
US-60-094-436-12

```

```

alignment_scores:
Quality: 2806.50 Length: 660
Ratio: 4.717 Gaps: 6
Percent Similarity: 90.152 Percent Identity: 80.909
alignment_block:
US-09-674-824-2 x US-60-094-436-12
Align seg 1/1 to: US-60-094-436-12 from: 1 to: 2008
4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuA5gAlaAs 20
14 TCGGCGGTGGCGCGCGCTGCTCCTC..... 40
20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34
41 .CTCGCGCGCGCGCGCTGCGCGCGCGCTGCGCGCGCGCGCGCGCGCG 89
35 ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
90 GCGCGCTCCAGCGCGCTGCTGCGCGCGCGCTGCTGCGCGCGCGCGCG 139
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
140 GAGGGCGCGCGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCGCGCT 189
67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84
190 CGTGGCGCGCTTCTCGCG...CGCGCGCGCGCGCGCGCGCGCGCG 236
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
237 CATCGACGCGCGCGCGCTGCGCGCGCGCTGCGCGCGCGCGCGCGT 283
101 AspLeuLeuLeuGlyIleAlaGluAspSerIleAspSerIleIleVa 117
284 ...CTCGAACCTGAAGGATGCTGAAGGTTCCATCATCAACACAGTAGT 330
117 laAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
331 TGTGGCAAGTGACCAAGATTCTGAGATTGTTGGTTGGAAGGAGCAAGCTC 380
134 laAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
381 GAGCTAAAGTAACACAAAGCATTTGCTTTGTAACCGCGCGAAGCTTCTCT 430
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167
431 TATGCAAGTCTGGGGTCTAGGAGATGTTGTTGTTTTCATTGCCAGTTGC 480
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
481 TCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 530
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200
531 ATGTTACCTCCCTGATAGAATTATGCAATGTCATTTTACACAGAAACAC 580
201 IleLysIleProCysPheGlySerHisGluValThrPhePheHisGl 217
581 ATTCGATTCCATGCTTTGGCGGTGAACATGAAGTTACCTTCTTCATGA 630
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234
631 GTATAGAGATTACAGTTGACTGGTGTGTTGTTGATCATCCCTCATATCACA 680
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250
681 GACCTGGAATTTATATGAGATAAGTTTGGTCTTTTGGTGATAATCAG 730
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267

```



|||||  
731 TTCAGATACACACCTCTTGTGCTATGCTGATGCTGAGGCTCCTTTGATCCT 780  
267 ucluleuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284  
|||||  
781 TGAATTGGAGGATATATTTATGGACAGAATTGCATGTTGTTCTCAATG 830  
284 sPTPrHisAlaSerLeuValProValLeuLeuAlaLysTyrArgPro 300  
|||||  
831 ATTGGCATGCCAGTCTAGTGGCAGTCCCTTCTGTGTCACAAATATAGACCA 880  
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317  
|||||  
881 TATGGTGTATTAAAGACTCCCGCAGCATCTTGTGAATACATAATTTAGC 930  
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProP 334  
|||||  
931 ACATCAGGGTGTAGAGCCTGCAAGCACATATCTCGACCTGGGTTGCCAC 980  
334 roGluTrpTyrGlyAlaLeuGluTyrValPheProGluTrpAlaArgArg 350  
|||||  
981 CTGAATGGTATGGAGCTCTGGAGTGGGTATTCCTGAATGGGCGAGGAGG 1030  
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367  
|||||  
1031 CATGCCCTTGACAAAGGTGAGCGAGTAAATTTTGAAGGTGCAGTTGT 1080  
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384  
|||||  
1081 GACAGCAGATCGAATCGTGACTGTCTAGTAGAGGTATTTCGTGGGAGGTCA 1130  
384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400  
|||||  
1131 CAACTGCTGAAGGTGACAGGGCTCCATAGAGCTCTTAAGCTCCACAAAG 1180  
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417  
|||||  
1181 AGTGATTAACGAATTTGTAATGGAATTGACATTAATGATTGGAACCC 1230  
417 oThrThrAspLysCysLeuProHisThrHisTyrSerValAspAspLeuSerG 434  
|||||  
1231 TGGCAGACAAATGATCCCTGTCTATTCTGTTGATGACCTCTCTG 1280  
434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450  
|||||  
1281 GAAAGGCCAAATGTAAGGTGCATTCAGAAAGGAGCTGGGTTTACCTATA 1330  
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467  
|||||  
1331 AGGCCTGATGTTCTCTGATGGCTTTATTGGAGGTTGGATTATCAGAA 1380  
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484  
|||||  
1381 AGGCATGATCTCATTCAACTATCATACCATCTCATCGCGGAAGATG 1430  
484 aGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTyrPmet 500  
|||||  
1431 TTCAAATTTGTCATGCTTGGATCTGGTGACCCAGAGCTTGAAGATTTGGATG 1480  
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTyrValGlyPh 517  
|||||  
1481 AGATCTACAGAGTGCATCTTCAAGGATAAATTTCTGGATGGGTTGGATT 1530  
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534  
|||||  
1531 TAGTGTTCAGTTTCCACCCCAAACTGCGCGCTGCGATATATTTGTTAA 1580  
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550  
|||||  
1581 TGCCATCCAGATTGCAACCTTGTGGTCTCAATCAGCTATATGCTATGCAG 1630  
551 TyrGlyThrValProValIleHisGlyThrGlyLeuArgAspThrVa 567  
|||||

1631 TATGGCACAGTTCTCTGTTCATGCACACTGGGGCCCTTAGAGATACCGT 1680  
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTyrPA 584  
|||||  
1681 GGAGAATCTCAACCCCTTCGGTGAGATGGAGAGAGGTACAGGTGGG 1730  
584 laPheSerProLeuThrValAspLysMetLeuTyrAlaLeuArgThrAla 600  
|||||  
1731 CATTCGCCACCCCTAACACAGAAACATCTTGTGGACATTTGGCAACTGCA 1780  
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617  
|||||  
1781 ATATCTACATACAGGAAACAAAGTCTCTCTGGGAAGGCTAATGAAGCG 1830  
617 gGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSer 633  
|||||  
1831 AGCATGTCAAAAGACTTCACGTGGGACCATCGCGCTGAACAATACGAAC 1880  
634 ArgSerSerGlyProSerTrpThrAsnProThrSerCysArgArgG1 650  
|||||  
1881 AAATCTTCCAGTGGCCTTCATCGATCGACCTATGTCATGTAAAAAA.. 1928  
650 yLeuGlyArgSerLysCysGluSerPro 659  
|||||  
1929 .....AGGACCAAACTGGTGTCTCT 1949  
seq\_name: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq:US-09-402-254-52  
seq\_documentation\_block:  
; Sequence 52, Application US/09402254  
; GENERAL INFORMATION:  
; APPLICANT: Guan, Hanning  
; APPLICANT: Keeling, Peter L.  
; TITLE OF INVENTION: PLANT LIKE STARCHES AND THE METHOD OF MAKING THEM IN  
; TITLE OF INVENTION: HOSTS  
; FILE REFERENCE: 2461-52  
; CURRENT APPLICATION NUMBER: US/09/402,254  
; EARLIER FILING DATE: 1999-10-01  
; EARLIER APPLICATION NUMBER: PCT/US98/06660  
; EARLIER FILING DATE: 1998-04-03  
; EARLIER APPLICATION NUMBER: 60/042,939  
; EARLIER FILING DATE: 1997-04-04  
; NUMBER OF SEQ ID NOS: 77  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 52  
; LENGTH: 1749  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1749)  
US-09-402-254-52  
alignment\_scores:  
Quality: 2713.50 Length: 585  
Ratio: 4.916 Gaps: 3  
Percent Similarity: 94.359 Percent Identity: 86.325  
alignment\_block:  
US-09-674-824-2 x US-09-402-254-52  
Align seg 1/1 to: US-09-402-254-52 from: 1 to: 1749  
45 ValAlaGluLeuSerArgGlyProAlaAlaArgProAlaGlnGln1 61  
|||||  
4 GTCCGGAGCTGAGCAGGAGGGGCCCGCGCCGCGCTGCCACCGC 53  
61 nGlnLeuAlaProProLeuValProGlyPheLeuAlaProProProA 78  
|||||  
54 GCTGCTGGGCGCCCGCTGTCGCGGCTTCTCGCG...CGCGCGCG 100  
78 laProAlaGlnSerProAlaProThrGlnProProLeuProAspAlaGly 94



995	GTTATTCTGGGGAGGTCAACA	CTGCTGAAGTGGACAGGGCCTCAATGAG	1044
395	LeuLeuSerSerArgLysSerVal	LeuAsnGlyIleValAsnGlyIleAs	411
1045	CTCTTAAGCTCCAGAAAGAGT	CTATTAAACGGAATGTAAATGGAATGA	1094
411	pileAsnAspTrpAsnProThr	ThrAspLysCysLeuProHisHisTrpS	428
1095	CATTAAATGATTGGACCCCTG	CCACAGACAAATGTATCCCTCTGTCATTATT	1144
428	erValAspAspLeuSerGlyLys	AlaLysCysLysAlaGluLeuGlnLys	444
1145	CTGTTGATGACCTCTCTGSA	AGGCCAAATGTAAAGGTGCATTGCAGAG	1194
445	GluLeuGlyLeuProValArg	GluaspValProLeuIleGlyPheIleG	461
1195	GAGCTGGGTTTACCTATAAG	CCCTGATCTCTCTGATTGGCTTTATTGG	1244
461	YargLeuAspTyrGlnLysGly	IleAspLeuLysMetAlaIleProG	478
1245	AAGGTGGATTATCAGAAAGC	ATTGATCTCATTTCAACTTATCATACAG	1294
478	luLeuMetArgGluaspValG	lnPheValMetLeuGlySerGlyAspPro	494
1295	ATTCTCATCGGGAAGATCT	CAATTTGTCATGCTTGGATCTGGTGACCCA	1344
495	IlePheGluGlyTrpMetArg	SerThrGluSerSerTyrLysaspLysPh	511
1345	GAGCTTGAAGATTGGATGAG	ATCTACAGTCTGACAGTCTGATCTCAAGGATAAATT	1394
511	eArgGlyTrpValGlyPheSer	ValProValSerHisArgIleThrAlaG	528
1395	TGCTGGATGGTGGATTAGT	GTGTTCCAGTTTCCACCGGAATRACTGCGG	1444
528	lyCysAspIleLeuLeuMetPro	SerArgPheGluProCysGlyLeuAsn	544
1445	GCTGGCATATATTGTTAATG	CCATCCAGATTCGAACCTTGTGGTCTCAAT	1494
545	GlnLeuTyrAlaMetGlnTyr	GlyThrValProValValHisGlyThrG	561
1495	CAGCTATATGCTATGCAGT	ATGGCACAGTTCCTGTTGTCCTCACTGCACTGG	1544
561	yGlyLeuArgaspThrValG	luThrPheAsnProPheGlyAlaLysGlyG	578
1545	GGGCTTTAGATACCGTGGAG	AATCTCAACCCCTTTCGTTGAGAAATGGAG	1594
578	luGluGlyThrGlyTrpAla	PheSerProLeuThrValaspLysMetLe	594
1595	AGCAGGTACAGGTGGGCAT	TTCCACCCCTTACCCACAGAAAACATGTTT	1644
594	uTrpAlaLeuArgThrAla	MetSerThrPheArgGluHisLysProSerT	611
1645	GTGGACATTGGAACTGCAAT	ATCATATACAGGGAACACAAGTCTCCTCCT	1694
611	rpGluGlyLeuMetLysArgG	lyMetThrLysaspHisThrTrpaspHis	627
1695	GGGAAGGCTTAATGAAGCG	AGGATGTCAAAAGACTTCACGTGGGACCAT	1744
628	Ala	628	
1745	GCC	1747	

```
seq_name: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:US-09-625-406-12
seq_documentation_block:
: Sequence 12, Application US/09625406
: GENERAL INFORMATION:
: APPLICANT: Keeling, Peter
: APPLICANT: Guan, Haining
: TITLE OF INVENTION: Starch Encapsulation
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Greenlee, Winner and Sullivan, P C
```



```

; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/625,406
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/941,445
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1752 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1752
; US-09-625-406-12

alignment_scores:
  Quality: 2713.50      Length: 585
  Ratio: 4.916          Gaps: 3
  Percent Similarity: 94.359  Percent Identity: 86.325

alignment_block:
US-09-674-824-2 x US-09-625-406-12 ..

Align seg 1/1 to: US-09-625-406-12 from: 1 to: 1752

45 ValAlaGluLeuSerArgGluGlyProAlaAlaArgProAlaGlnGlnG1 61
|||||
4 GTCCGGAGCTGACGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 53
|||||

61 nGlnLeuAlaProProLeuValProGlyPheLeuAlaProProProA 78
|||||
54 GCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 100
|||||

78 laProAlaGlnSerProAlaProThrGlnProProLeuProAlaGly 94
|||||
101 AGCCACAGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 150
|||||

95 valGlyGluLeuAlaProAlaProAlaGluGlyIleAlaGluAspSe 111
|||||
151 CTGGGGGACCTCGT.....CTCGAACCTGAAGGATTGCTGAAGGTTTC 194
|||||

111 rleAspSerIleIleValAlaAlaSerGluGlnAspSerGluIleMeTa 128
|||||
195 CATCGATAACACAGTAGTGTGGCAAGTGGCAAGATTCTGAGATTGTGG 244
|||||

128 spAlaAsnGluProGlnAlaLysValThrArgSerIleValPheVal 144
|||||
245 TTGGAAGGAGCAAGCTCGAGCTAAAGTACACAAAGCATTGTCTTTGTA 294
|||||
```



```
445 GluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheIleGl 461
|||||
1195 GAGCTGGGTTTACCTATAGGCGCTGATGTTCTCTGATGGTTTATGG 1244
|||||
461 YArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIleProG 478
|||||
1245 AAGGTTGGATTATCAGAAGGCAATGATCTCATTTCAACTTATCATACCAG 1294
|||||
478 LuLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAspPro 494
|||||
1295 ATCTCATCGGGAAGATGTTCAATTTGTCAAGTCTGGATCTGGTCAACCA 1344
|||||
495 IlePheGluGlyTyrMetArgSerThrGluSerSerTyrLysAspLysPh 511
|||||
1345 GAGCTTGAAGATTGGATGAGATCTACAGATCGATCTTCAAGGATAAAT 1394
|||||
511 eArgGlyTyrValGlyPheSerValProValSerHisArgIleThrAlaG 528
|||||
1395 TCGTGGATGGGTTGGATTAGTGTTCAGTTTCCACCGAATAACTGCCG 1444
|||||
528 LysAspIleLeuLeuMetProSerArgPheGluProCysGlyLeuAsn 544
|||||
1445 GCTGGATATATGTTATGCAATCCACATTCGAACCTTGGTCTCAAT 1494
|||||
545 GlnLeuTyrAlaMetGlnTyrGlyThrValProValValHisGlyThrGl 561
|||||
1495 CAGCTATATGTCAGTATGCGACAGTTCCTGTTGCCATGCAACTGG 1544
|||||
561 yGlyLeuArgAspThrValGluThrPheAsnPropPheGlyAlaLysGlyG 578
|||||
1545 GGGCTTAGACATACCGTGGAGAACTTCAACCTTTCGGTGAGAAATGGAG 1594
|||||
578 LuGlyGlyThrGlyTyrAlaPheSerProLeuThrValAspLysMetLe 594
|||||
1595 AGCAGGTACAGGTGGGCATTCGACCCCTAACACAGAAACATGTTT 1644
|||||
594 uTyrAlaLeuArgThrAlaMetSerThrPheArgGluHisLysProSerT 611
|||||
1645 GTGACATGCGAACTGCAATATCTACATACAGGGAACACAAAGTCCCT 1694
|||||
611 rpGlyGlyLeuMetLysArgGlyMetThrLysAspHisThrTyrAspHis 627
|||||
1695 GGGAAAGGCTAATCAAGCGAGCATGTCAAAGACTTCAAAGACTTCACG 1744
|||||
628 Ala 628
|||
1745 GCC 1747
```

seq\_name: /cgn2\_6/ptodata/1/pna/US090\_COMB.seq:US-09-077-564-1

```
seq_documentation_block:
; Sequence 1, Application US/09077564
; GENERAL INFORMATION:
; APPLICANT: Knight, Mary E.
; APPLICANT: Keeling, Peter L.
; TITLE OF INVENTION: Modification of Starch Synthesis in
; TITLE OF INVENTION: Plants
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Ag Products
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,564
; FILING DATE: 14-DEC-1998
```

```
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/02990
; FILING DATE: 04-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9524938.9
; FILING DATE: 06-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: SEE 45052/UST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-1699
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2992 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: NUMBER 1
; US-09-077-564-1

alignment_scores:
Quality: 2655.50 Length: 745
Ratio: 4.304 Gaps: 11
Percent Similarity: 82.819 Percent Identity: 71.544

alignment_block:
US-09-674-824-2 x US-09-077-564-1 ..
Align seg 1/1 to: US-09-077-564-1 from: 1 to: 2992

4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuArgAlaAs 20
|||||
614 TCGCGCGTGGCGCGCGGTGCTC.....CTCCTCGCGCG 648

20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAlaLeu 37
|||||
649 GNGCGCTGGCGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGTCC 698

37 rArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArgGluGlyPro 53
|||||
699 AGCGGTGCTGCGCGCGCGGTGCTGCGGAGCTGACGAGGAGGCGCC 748

54 AlaAlaArgProAlaGlnGlnGlnGlnLeuAlaProProLeuValProGl 70
|||||
749 GCGCGCGCGCGCGTGGCACCGCGCGTCTGGCGCGCGCGCGTGGCGCG 798

70 yPheLeuAlaProProProAlaProAlaGlnSerProAlaProThrG 87
|||||
799 CTTCCTCGCG...CCGCGCGCGCGCGCGCGCGGTGAGCGCGCATCGACG 845

87 InProProLeuProAspAlaGlyValGlyGluLeuAlaProAspLeu 103
|||||
846 CGCGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGT.....CTCGAA 889

104 LeuGlyGlyIleAlaGluAspSerIleValAlaAla 120
|||||
890 CTGAAGGAGATTGCTGAAGGTTCCATCGATAACACAGTAGTTGTGGCAAG 939

120 rGluGlnAspSerGluIleMetAspAlaAsnGlnProGlnAlaLysV 137
|||||
940 TGAGCAAGATTCTGAGATTGTGGTGAAGAGGAGCAAGCTCGAGCTAAG 989

137 alThrArgSerIleValPheValThrGlyGluAlaAlaPro.TyrAla 153
|||||
990 TAACACAAAGCATTTGCTTTTGAACCGCGCGAGCTTCTCTTAATCGAA 1039

153 sSerGlyGlyLeuGlyValCysGlySerLeuProIleAlaLeuAla 170
|||||
```



1040	CTCTGGGGGCTTAGAGAGATCTTTGTGGTTTCATTGGCCAGTTGCTCTTGTCTG	1089
170	laArgGlyHisArgValMetValValMetProArgTyrLeuAsnGlySer	186
1090	CTCGGGGTCAACCGTGTATGGTTGTAAATGCCAGACATTAAATGGTACC	1139
187	SerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLysAl	203
1140	TCCGATAGAATAATGCAATGCATTTTACTCAGAAAAACACATTCGGAT	1189
203	eProCysPheGlyGlySerHisGluValThrPhePheHisGluTyrArg	220
1190	TCCATCTTTGGCGGTGAACATGAAGTACCTTCTTCCATGAGTATAGAG	1239
220	sPasnValAspTrpValPheValAspHisProSerTyrHisArgProGly	236
1240	ATTCAATTCAGTGGGTGTGTGTGATCATCCCTCATATCACAGACCTGGA	1289
237	SerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyr	253
1290	AATTTATATGAGATAGTTTGGTGGCTTTTGGTGATTAATCAGTTACAGATA	1339
253	rThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuG	270
1340	CACACTCTTTGCTATGCTGCATGTAGGCTCTTTGGTCTTGAATTTGG	1389
270	lyGlyTyrIleTyrGlyClnAsnCysMetPheValValAsnAspTrpHis	286
1390	GAGGATATATTATGACAGAAATTCGATGTGGTTGGTTGTCAAATGATGGCAT	1439
287	AlaSerLeuValProValLeuLeuAlaAlaLysTyrArgProTyrGlyVal	303
1440	GCCAGTCTAGAGCCAGTCTTCTTGGCTGCAAAATATAGACCATATGGTGT	1489
303	lTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisGlnG	320
1490	TTATAAAGACTCCCGACGATCTGTGTAATACATAAATTTAGCACATCAGG	1539
320	lyValGluProAlaSerThrTyrProAspLeuGlyLeuProGluTyr	336
1540	GTGTAGAGCTGCAGACACATATCTGNACCTTGGTGGTCCACCTGAATGG	1589
337	TyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAlaLe	353
1590	TATGGAGCTCTGAGTGGGTATTCCTCAATGGCGGAGGAGCATGCGCT	1639
353	uasLysGlyGluAlaValAsnPheLeuLysGlyAlaValValThrAlaA	370
1640	TGACAAAGGTGAGGCAGTAAATTTTGTAAAGGTGCAGTTGTGTGACACAG	1689
370	sPArgIleValThrValSerGlnGlyTyrSerTrpGluValThrThrAla	386
1690	ATCGAATCTGTACTGTAGTAAGGGTTATTCATGGAGGTGCACAACCTGT	1739
387	GluGlyGlyClnGlyLeuAsnGlnLeuLeuSerSerArgLysSerValLe	403
1740	GAAGGTGGACAGGGCTCAATGAGCTCTTAAGCTCCAGAAAGAGTGTATT	1789
403	uasnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrThrA	420
1790	AAACGGAATTTGTAATGGAAATGACATTAAATGATTGGAACCTCCACACAG	1839
420	sPlyCysLeuProHisHisTyrSerValAspAspLeuSerGlyLysAla	436
1840	ACAAATGTATCCCTCTCATATTATCTGTGTGATGACCTCTCTTGAAGGCT	1889
437	LysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAs	453
1890	AAATGTAAAGGTGCATTGCAAGAGAGGCTGGGTTTACCTATAAGGCCTGA	1939
453	pValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleA	470
1940	TGTTCTCTGATTGGCTTTATTGGAAGATTGGATTATCAGAAGGCATTG	1989

470 spLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGlnPhe 486  
1990 ATCTCATTCAACTTATCATACCAAGATCTCATGCGGAAGAAATGTTCAATTTT  
487 ValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgSerTh 503  
2040 GTCATGCTTGGATCTGGTGACCCAGACCTTCAAGATGGATGAGATCTAC 2089  
503 tGluSerSerTrpLysAspLysPheArgGlyTrpValGlyPheSerValP 520  
2090 AGAGTCGATCTTCAAGGATAAATTTCTGGATGGGTGGATTAGTGTTC 2139  
520 toValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSer 536  
2140 CAGTTTCCACCGAATAACTCGCGCTCGCATATATGTTAATGCCATCC 2189  
537 ArgPheGluProCysGlyLeuAsnGlnLeuTrpAlaMetGlnTrpGlyTh 553  
2190 AGATTTCGAACCTTGTGGTCTCAATCAGCTATATGCTATGCAGTATGGCAC 2239  
553 rValProValValHisGlyThrGlyGlyLeuArgAspThrValGluThrP 570  
2240 AGTTCCTGTGTCCATGCACACTGGGGCCCTAGAGATACCGTGGAGAACT 2289  
570 heAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpAlaPheSer 586  
2290 TCAACCTTTTCGGTCAGAAATGGAGAGCAGGGTACAGGGTGGGCAATTCGCA 2339  
587 ProLeuThrValAspLysMet\_LeuTrpAlaLeuArgThrAlaMetSerT 603  
2340 CCCCATACCACAGAAACATGTTGTGGCATTGCGAACTGCAATATCTA 2389  
603 hrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMet 619  
2390 CATACAGGAACACACAAGTAATAATGGGAAGGCCTAATGAAGCCAGCATG 2439  
620 ThrLysAspHisThrTrpAspHisAlaPro\_SerSerThrSerArgSerS 636  
2440 TCAAAAGAGTTCACGTGGAGCACCCTGCCCTGAAACAATACAGACAAATCTT 2489  
636 exSerGlyProSerTrpThrAsnProThrSerCysArg\_ArgGlyLeuGlu 652  
2490 CCAGTGGGCTTCATCGGATCGACCC...GATGTTCAATGGAAAAAGGG 2536  
652 yArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSerSerP 669  
2537 ACCAAAGTTGGTTCTCTTCAAGATTATCAGTTTCATCTATAGTA 2586  
669 heArgGlyProGluGlyTrpPro....CysThrLeuArgCysProAla 683  
2587 AGCTGAATGATGAAGAAACCCCTGTACATTACATGAAGGCAGACCGG 2636  
684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPhe..... 695  
2637 CTATTGGCTCCATTGCTCAATGTCGTGTTGGCTGCCTTCGCTCGATGG 2686  
695 ..... 695  
2687 ACCGGATGCAGTGAGGAATCCAGNCGAAGCAGCTTTTGAAGGATAGGAA 2736  
696 .....AlaGlySerArgThrTrpAspGlyCysAlaAlaAlaValThr 710  
2737 GGGAGCTTGAAGC...AGTCAGCGAGCAGCGCAAGCCCTTCGCCGTTAAT 2783  
711 AlaSerGlyGlyArgGlnLeuGlnPhe 719  
2784 TCATATGGAAACAAGCTGGAGTCAGTTT 2810

```
seq_name: /cgn2_6/ptodata/1/pna/US096C_COMB.seq: US-09-654-617-4517533
seq_documentation_block:
: Sequence 451753, Application US/09654617
```



## ; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; TITLE OF INVENTION: Annotated Plant Genes  
; FILE REFERENCE: 38-21(15097)D  
; CURRENT APPLICATION NUMBER: US/09/654.617  
; CURRENT FILING DATE: 2000-09-05  
; NUMBER OF SEQ ID NOS: 463173  
; SEQ ID NO 451753  
; LENGTH: 2216  
; TYPE: DNA  
; ORGANISM: Sorghum bicolor  
US-09-654-617-451753

alignment\_scores:  
Quality: 2578.00 Length: 628  
Ratio: 4.628 Gaps: 9  
Percent Similarity: 88.694 Percent Identity: 78.662  
alignment\_block:  
US-09-674-824-2 x US-09-654-617-451753  
Align seg 1/1 to: US-09-654-617-451753 from: 1 to: 2216

107 lleAlaGluAspSerIleAspSerIleValAlaAlaSerGluGlnAs 123  
1 ATGTGTAAGGTTCATCGATGAGACAGTAGTTGTGGCAAGTGGACAGA 50  
123 pSerGluIleMetaspAlaAsnGluGlnProGlnAlaLysValThrArgS 140  
51 TTCGAGATCGGTGGTGAAGGAGCAAGCTCGAGCTAAAGTAAACAAA 100  
140 erIleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGly 156  
101 GCATTGTCTTTGTAACCTGGCGAAGCTTCTCTTATGCAAGCTCGGGGT 150  
157 LeuGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyH1 173  
151 CTAGAGATGTTGTGGTTCATTCGCAAGTGTCTTCTGCTGCTGCTCA 200  
173 sArgValMetValValMetProArgTyrLeuAsnGlySerSerAspLysA 190  
201 CCGTGTGATGTTGTAATGCCACAGATTTAAATGGTACCTCTGTATAGA 250  
190 snTyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPhe 206  
251 ATTACGCAATGTCATTTTACACAGAAAGCACATTCGGAATCCATGCTT 300  
207 GlyGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValas 223  
301 GCGGTGAACATGAAGTACCTTTTCCATGAGTACAGATTCAGTTGA 350  
223 pTrpValPheValAspHisProSerTyrHisArgProGlySerLeuTyrG 240  
351 CTGGGTGTTTGTGATCATCCCTCATATCACACACCTGGAAATTTATATG 400  
240 lYAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeu 256  
401 GAGATAAGTTGGTGTCTTTGGTGAATATCAGTTTCAGATACAGCTCCTT 450  
257 CysTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyTyrIle 273  
451 TGCTATGCTGCATGTGAGGCTCTTTGGTCTTGAATTCGGAGGATATAT 500  
273 eTyrGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuV 290  
501 TTTATGACAGAATTCGATGTTTGTGTAATGATTGCCATGCCAGCTAG 550  
290 alProValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAsp 306  
551 TGCCAGTCTCTTCTGCTGCAAAATATAGACCATATGTTGTTTATAAAGAC 600

307 SerArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluPr 323  
601 TCCCGCAGCATCTCTGTANTACATAATTTAGCACATCAGGTGTAGAGCC 650  
323 oAlaSerThrTyrProAspLeuGlyLeuProProGluTyrTyrGlyAlaL 340  
651 TSCAAGCACATATCTACCTTGGTTGCCACCTGAATGGTATGAGGCTC 700  
340 euGluTyrValPheProGluTrpAlaArgArgHisAlaLeuAspLysGly 356  
701 TGGAGTGGGTATTCCTGAATGGCAAGGAGGCATGCCCTTGACAGGGT 750  
357 GluAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleVa 373  
751 GAGCAGTAAATTTTGAAGGTGCAGTTGTGACACGACATCGAATGT 800  
373 lThrValSerGlnGlyTyrSerTrpGluValThrThrAlaGluGlyGly 390  
801 GACTGTGAGTAAGGTTTATTCTGGGAGGTCAACAACCTGCTGAAGGTGAC 850  
390 lnGlyLeuAsnGluLeuSerSerArgLysSerValLeuAsnGlyIle 406  
851 AGGCTCTCAATGAGCTCTTAAGCTCCGGAAGAGTGTATTAAACGGANT 900  
407 ValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLe 423  
901 GTAATGGAATTCACATTAATGATTGGACCTCGACGCGACAAATGTAT 950  
423 uProHisHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysA 440  
951 CCCTTGTCTATTCTGTTGATGACCTCTCTGAAAAGCCAAATGTAAAA 1000  
440 laGluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeu 456  
1001 GTCCATTCGAGAAGGAGCTGGGTTTACCTATAAGGCTGAAGTTCCTCTG 1050  
457 IleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleY 473  
1051 ATTTGGTTTATTGGAAGATTGGATTATCAGAAAGGCAATGTCTCATTC 1100  
473 sMetAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuG 490  
1101 ACTTATCATACCATCTCATCGGGACGACGCTTCAATTTGTCATGCTTG 1150  
490 lYSerGlyAspProIlePheGluGlyTyrPheMetArgSerThrGluSerSer 506  
1151 GATCTGCTGACCCAGCTCGAAGACTGGATGAGATCTACAGATCGGAC 1200  
507 TyrLysAspLysPheArgGlyTyrValGlyPheSerValProValSerHi 523  
1201 TTCAAGGATAAATTTCTGATGGGTTGGATTAGTTTCCAGTTTCCCA 1250  
523 sArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluP 540  
1251 CGGAATAACTGCGGCTGCGATATATTGTTAATGCCATCCAGATTCGAAC 1300  
540 roCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProVal 556  
1301 CTGTGGTCTCAATCAGCTATATGATGTCATATGCGACATGCGACATCTCTGT 1350  
557 ValHisGlyThrGlyLeuArgAspThrValGluThrPheAsnProPh 573  
1351 GTCCATGCAACTGGGGGCTTAGAGATACTGTGGAGAACTTCAACCCCTT 1400  
573 eGlyAlaLysGlyGluGlyThrGlyTyrPheAlaPheSerProLeuThrV 590  
1401 CGCTCAGAAATGGAGAACAGGGTACAGGTGGGCAATTCGACCCCTTAA 1450  
590 alAspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgG 606  
1451 CGGAAAACATGTTCTGGACATTCGCAACTGCAATTTCCACATACAGGA 1500  
606 uHisLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspH 623



```
1501 GCACAAATCTCTCGGAAGGGCTCATGAGAGGGGATGCTCAAAAGACT 1550
623 isThrTrpAspHisAlaPro..SerSerThrSerArgSerSerGlyPr 639
1551 TCACGTGGACCATGCCGCTGAAACAAATACGAACAAATCTTCCAGTGGGC 1600
639 oSerTrpThrAsnProThrSerCysArgArgGlyLeuGlyArgSerLysC 656
1601 TTTCATCGATCGACCTGATGTCATGTAAGGGGAC..... 1634
656 ysGlu.SerProSerAlaLeuLysThr...SerSerSerSerPheArgG 671
1635 ..CACAAGTGTGTCTTAAAGATCTTCAATCTTCACTACATGG 1682
671 yProGlu.GlyTyProCysThrLeuArgCysProAlaThrValGluSer 687
1683 AAGGATGAGAAACCCCTGTATACAT..... 1711
688 GlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrTrpAspGlyCy 704
1712 .....GGAAGCGGACC.....GGCTG 1728
704 saAlaAlaAlaAlaValThrAlaSerGlyGlyArgGlnLeuGlnPheTrpG 721
1729 TTGGGTGGCTCCATGTCCAGTTAGCCTTACGGAGCCTGCTGTTGGC 1778
721 lyLeuArgLysGlyCysAlaAlaGly 729
1779 TGGTT.....TGTCACCAAGGA 1795
```

seq\_name: /cgn2\_6/ptodata/1/pna/US096E\_COMB.seq:US-09-684-016-451753

```
seq_documentation_block:
; Sequence 451753, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 451753
; LENGTH: 2216
; TYPE: DNA
; ORGANISM: Sorghum bicolor
US-09-684-016-451753
```

alignment\_scores:  
Quality: 2578.00 Length: 628  
Ratio: 4.628 Caps: 9  
Percent Similarity: 88.694 Percent Identity: 78.662

alignment\_block:  
US-09-674-824-2 x US-09-684-016-451753 ..  
  
Align seg 1/1 to: US-09-684-016-451753 from: 1 to: 2216  
  
107 ileAlaGluAspSerIleAspSerIleValAlaAlaSerGluClnAs 123  
1 ATTGCTGAAGGTTCCATCGATGACAGAGTAGTTGTGGCAAGTGAGCAAGA 50  
  
123 pSerGluIleMetAspAlaAsnGluGlnProGlnAlaLysValThrArgS 140  
51 TTCTGAGATCGTGTGGGAAGGAGCAAGCTCCAGCTAAAGTAACTAACACA 100  
  
140 erIleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGly 156  
101 GCATTGCTTTTAACCTGGCGAAGCTCTCTTATGCAAAAGTCTGGGGGT 150

```
157 LeuGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHI 173
151 CTAGGAGATGTTGGTTCATTCAGTTGCTTGTGCTCGTGGTCA 200
173 sArgValMetValValMetProArgTyrLeuAsnGlySerSerAspLysA 190
201 CGGTGTGATGGTGTAAATCCAGATATTTAAATGGTACCTCTGATAGA 250
190 snTyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPhe 206
251 ATTACCAAAATGCATTTTACAGAAAACACATTCGGATTCCATGCTTT 300
207 GlyGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValAs 223
301 GGCGGTGAACATGAAGTTACCTTTTCCATAGTACAGAGATTTCAGTTGA 350
223 pTrpValPheValAspHisProSerTyrHisArgProGlySerLeuTyrG 240
351 CTGGGTGTTTGTGATCATCCCTCATACAGACCTGGAAATTTATATAG 400
240 lyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeu 256
401 GAGATAAGTTTGGTCTTTTGGTGATAAATCAGTTTCAGATACACGCTCT 450
257 CysTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIL 273
451 TGCTATGCTGCATGTCAGGCTCCTTTGGTCTTGAATTTGGGAGGATATAT 500
273 eTyrGlyGlnAsnCysMetPheValValAsnAspTrpHisAlaSerLeuV 290
501 TTATGGACAGAAATTCATGTTTGTGTGAATGATGGCATGCCAGCTCTAG 550
290 alProValLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAsp 306
551 TGCAGTCTTCTTGTCTGCAAAATATAGACCATATGGTGTATAAGAC 600
307 SerArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluPr 323
601 TCCCGCAGCATTCCTGTAATACATAATTTAGCACATCAGGGGTGAGAGCC 650
323 oAlaSerThrTyrProAspLeuGlyLeuProProGluTyrTrpTyrGlyAla 340
651 TGCAGGACACATATCTGCACCTTGGGTTGCCACCTGAAATGTATGGAGCTC 700
340 euGluTrpValPheProGluTrpAlaArgArgHisAlaLeuAspLysGly 356
701 TGGAGTGGGTATTCCTGAAATGGCAAGAGGCATGCCCTTGACAAGGGT 750
357 GluAlaValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleVa 373
751 GAGGCGATTAAATTTTGAAGGTCGAGTTGTGACAGCAGATCCGAATGT 800
373 lThrValSerGlnGlyTyrSerTrpGluValThrThrAlaGluGlyGlyG 390
801 GACTGTGATGAAGGTTATTCATGGAGGTCACAACCTGCTGAAGGTGGAC 850
390 lngLysLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIle 406
851 AGGGTCTCAATCAGCTCTTAAGCTCCCCGAAAGAGTGTATTAAACGGAAT 900
407 ValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLe 423
901 GTAATGAATGACATTAATGATTGGAACCTTCGACGGACAAATGTAT 950
423 uProHisHisTyrSerValAspAspLeuSerGlyLysAlaLysCysLysA 440
951 CCGTGTGTCATTTATCTGTGTGATGACCTCTCGAAAAGCCAAATGTAAAA 1000
440 laGluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeu 456
1001 GTGCATTGCAGAGAGGCTGGGTTTACCTATATAAGGCCTGAAGTTCTCTG 1050
```



```

457 ILeGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleY 473
|||||
1051 ATTGGTTTATTGGAAGATTGGATTATCAGAAAGGCAATGATCTCATTC 1100
473 sMetAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuG 490
|||||
1101 ACTTATCATACCATCTCTATCGGACGACGTTCAATTTGTCTGCTTG 1150
490 lySerGlyAspProIlePheGluGlyTyrMetArgSerThrGluSerSer 506
|||||
1151 GATCTGCTGACCCAGACCTCGAAGACTGGATGAGATCTACAGATCGGAC 1200
507 TyrLysAspLysPheArgGlyTrpValGlyPheSerValProValSerHi 523
1201 TTCAGGATAAATTTCTGGATGGGTTGATTTAGTTTCCAGTTTCCCA 1250
523 sArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluP 540
|||||
1251 CCGAATACTCGCGCTGCGATATATGTTAATGCCATCCAGATTCCGAAC 1300
540 roCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProVal 556
1301 CTTGTGCTCAATCAGCTATATGCTATGTCAGTATGGCACAGTTCTCT 1350
557 ValHisGlyThrGlyLeuArgAspThrValGluThrPheAsnProPh 573
1351 GTCCATGCAACTGGGGGCTTAGAGATACTGTGGAGAACTTCACCCCTT 1400
573 eGlyAlaLysGlyGluGlyThrGlyTyrAlaPheSerProLeuThrV 590
1401 CGGTGAGATGAGAACAGGTACAGGTGGGATTCGCACCCCTTAACCA 1450
590 aLaspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgG 606
1451 CGGAATACTGTTGTGGACATTCGAACTGCAATTTCCACATACAAGGA 1500
606 uHisLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAsp 623
1501 GCACAAATCTCTCTGGGAGGCTCATGAAGGGGGCATGTCAAAAGACT 1550
623 IsThrTrpAspHisAlaProSerSerThrSerArgSerSerGlyPr 639
1551 TCACGTGGGACCATGCCCTGAACAATACGAACAATCTTCAGTGGGCC 1600
639 oSerTrpThrAsnProThrSerCysArgArgGlyLeuGlyArgSerLys 656
1601 TTCATCGATCGACCCCTATGTCATGTAAAGGGGAC..... 1634
656 ysGluSerProSerAlaLeuLysThr...SerSerSerPheArgG 671
1635 ..CACAGTGTGTCTCTTAAGATCTTCAGTTCTTCATCCTACAGTGG 1682
671 yProGluGlyTyrProCysThrLeuArgCysProAlaThrValGluSer 687
1683 AAGGATGAGAAACCCCTGTACATTACAT..... 1711
688 GlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrTyrAspGlyCy 704
1712 .....GGAAGGCGGAC.....GGCTG 1728
704 sAlaAlaAlaAlaValThrAlaSerGlyGlyArgGlnLeuGlnPheTrp 721
1729 TTGGTGGCTCCAATGTCCAGGTAGCTTACGGAGCTCTGCTGCTTGGC 1778
721 lyIleArgLysGlyCysAlaAlaGly 729
1779 TGGTT.....TGTCCTCAAGGA 1795

```

```

seq_name: /cgn_6/ptodata/1/pna/US096B_COMB.seq:US-09-625-406-20
seq_documentation_block:
; Sequence 20, Application US/09625406
; GENERAL INFORMATION:

```

```

; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/625,406
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/941,445
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1620 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1620
; US-09-625-406-20

```

```

alignment_scores:
Quality: 2546.50      Length: 538
Ratio: 4.935          Gaps: 3
Percent Similarity: 95.911 Percent Identity: 87.918

alignment_block:
US-09-674-824-2 x US-09-625-406-20 ..

Align seg 1/1 to: US-09-625-406-20 from: 1 to: 1620

95 ValGlyGluLeuAla...ProAspLeuLeuLeu.....GluGlyIleAl 108
|||||
4 GTCCGGAGCTGAGCAGGAGGACCTCGGTCTCGAACCTGAAGGGATTGC 53

108 aGluAspSerIleAspSerIleIleValAlaAlaSerGluGlnAspSerG 125
|||||
54 TGAAGTTTCATCGATAACACACAGTAGTTGTGGCAAGTGAGCAAGATTCTG 103

125 luIleMetAspAlaAsnGluGlnProGlnAlaLysValThrArgSerIle 141
|||||
104 AGATTGTGGTTGCAAGGAGCAGCTCGAGCTAAAGTAACACAGCAATT 153

142 ValPheValThrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuG 158
|||||
154 GTCTTTGTAAACCGCGAAGCTTCTCTCTTATGCAAGCTTGGGGTCTAGG 203

158 yAspValCysGlySerLeuProIleAlaAlaAlaArgGlyHisArgV 175
|||||
204 AGATGTTTGTGGTTTCATGCCAGTTGCTTCTGCTGCTGTCACCGTG 253

```



175 alMetValValMetProArgTyrLeuAsnGlySerSerAspLysAsnTyr 191  
|||||  
254 TGATGGTGTAAATGCCAGATATTTAAATGGTACCTCCGATAAAGATAT 303  
|||||  
192 AlAlysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheGlyG 208  
|||||  
304 GCAATGCAATTTACACAGAAACACATTCGGATTCCATGCTTTGGGG 353  
|||||  
208 yserHisGluValThrPhePheHisGluTyrArgAspAsnValAspTrpv 225  
|||||  
354 TGAACATCAAGTACCTTCTCCATGAGTATAGAGATTCACTGACTGGG 403  
|||||  
225 alPheValAspHisProSerTyrHisArgProGlySerLeuTyrGlyVal 241  
|||||  
404 TGTGTGTGATCATCCCTCATATCACAGCCTGGAAATTTATATGGAGAT 453  
|||||  
242 AsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCysTy 258  
|||||  
454 AAGTTTGGTGTCTTTGGTGATTAATCAGTTTCAGATACACACTCCTTTGCTA 503  
|||||  
258 rAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrG 275  
|||||  
504 TGCTGCATGTGAGGCTCCTTTGATCCTTGAATTTGGGAGGATATATTTATG 553  
|||||  
275 LyGlnAsnCysMetPheValValAsnAspTyrPheHisAlaSerLeuValPro 291  
|||||  
554 GACAAATTCATGTTGTTGTCATGATTTGGCATGCCAGTCTAGTGCCA 603  
|||||  
292 ValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSerAr 308  
|||||  
604 GTCTTCTTGTGCAAAATATAGACCATATGTTGTTTATTAAGACTCCCG 653  
|||||  
308 gSerThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProAlas 325  
|||||  
654 CAGCATCTCTGTAATACATAATTTAGCACATCAGGGTGTAGAGCCTGCA 703  
|||||  
325 exThrTyrProAspLeuGlyLeuProProGluTyrTyrGlyAlaLeuGlu 341  
|||||  
704 GCACATATCTGACCTTGGTGGTGGCCACCTGAATGTTATGGAGCTGGAG 753  
|||||  
342 TrpValPheProGluTrpAlaArgHisAlaLeuAspLysGlyGluAl 358  
|||||  
754 TGGGTATTCCTTGAATGGCGAGGAGCATGCCCTTGACAAAGGAGGAGC 803  
|||||  
358 aValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValThrV 375  
|||||  
804 AGTTAATTTTGAAGGTGCAGTTGTGACAGCAGATCGAATCGTGACTG 853  
|||||  
375 alSerGlnGlyTyrSerTrpGluValThrAlaGluGlyGlnGly 391  
|||||  
854 TCAGTAAGGGTTATTCGTGGAGGTCAACACTGCTGAAGGTGGACAGGC 903  
|||||  
392 LeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleValAs 408  
|||||  
904 CTCATGAGCTCTTAAGCTCCAGAAAGATGTTAATAACGGAATGTATA 953  
|||||  
408 nGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuProH 425  
|||||  
954 TGAATTTGACATTAATGATTGGAAACCTGCCACAGACAAATGTATCCCT 1003  
|||||  
425 lshIstYrSerValAspAspLeuSerGlyLysAlaLysCysLysAlaGlu 441  
|||||  
1004 GTCATTTATCTGTGATGACCTCTCGAAAGGCCAAATGTAAGGTGCA 1053  
|||||  
442 LeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuIleG 458  
|||||  
1054 TTCAGAGAGAGGTGGGTTTACCTATTAAGGCTGATGTTCTCTGATTGG 1103  
|||||  
458 yPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetA 475  
|||||  
1104 CTTTATTGGAAGGTGGATTATCAGAAAGGCATTGATCTCATTTCAACTTA 1153  
|||||

475 latleProGluLeuMetArgGluAspValGlnPheValMetLeuGlySer 491  
|||||  
1154 TCATACCAGATCTCATCGGGAAGATGTTCAATTTGTCATGCTTGATCT 1203  
|||||  
492 GlyAspProIlePheGluGlyTyrMetArgSerThrGluSerSerTyrIly 508  
|||||  
1204 GGTGACCCAGAGCTTGAAGATTGGATGAGATCTACAGAGTCGATCTTCAA 1253  
|||||  
508 sAspLysPheArgGlyTyrValGlyPheSerValProValSerHisArgI 525  
|||||  
1254 GGATAAATTTCTGGTGGTGGTGGATTTAGTTTCCAGTTTCCACCGAA 1303  
|||||  
525 leThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluProCys 541  
|||||  
1304 TAACGCGCGCTCGATATATTTGTTAATGCCATCCAGATTTCGAACCTTGT 1353  
|||||  
542 GlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValValH 558  
|||||  
1354 GGTCTCAATCAGCTATATGCTATGCAGTATGGCACAGTTCTCTGTTGCCA 1403  
|||||  
558 sGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheGlyA 575  
|||||  
1404 TGCAACTGGGGCCTTAGAGATACCGTGAGAACTTCAACCTTTTCGGTG 1453  
|||||  
575 laLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrValasp 591  
|||||  
1454 AGAATGGAGAGCAGGGTACAGGGTGGGCATTCGCACCCCTAACCCAGAA 1503  
|||||  
592 LysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisL 608  
|||||  
1504 AACATGTTGGGACATTCGCACTGCAATATCTACATACAGGGAACACA 1553  
|||||  
608 ysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisThr 624  
|||||  
1554 AGTCTCTCTGGGAAGGCTAATGAAGCGAGGCATGTCAAAAGACTTCACG 1603  
|||||  
625 TrpAspHisAla 628  
|||||  
1604 TGGGACCATGCC 1615

seq\_name: /cgn2\_6/ptodata/1/pna/us093\_COMB.seq:us-09-345-214-6

seq\_documentation\_block:

; Sequence 6. Application US/09345214  
; GENERAL INFORMATION:  
; APPLICANT: Lightner, Jonathan E.  
; APPLICANT: Broglie, Karen E.  
; TITLE OF INVENTION: MODIFICATION OF STARCH BIOSYNTHETIC ENZYME GENE  
; TITLE OF INVENTION: EXPRESSION TO PRODUCE STARCHES IN GRAIN CROPS  
; FILE REFERENCE: BB-1147  
; CURRENT APPLICATION NUMBER: US/09/345,214  
; CURRENT FILING DATE: 1999-06-30  
; EARLIER APPLICATION NUMBER: 060/094,436  
; EARLIER FILING DATE: 1998-07-28  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 6  
; LENGTH: 1528  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-345-214-6

alignment\_scores:  
Quality: 2495.00 Length: 508  
Ratio: 5.051 Gaps: 1  
Percent Similarity: 97.244 Percent Identity: 89.567

alignment\_block:

US-09-674-824-2 x US-09-345-214-6/rev ..

Align seg 1/1 to reverse of: US-09-345-214-6 from: 1 to: 1528



134 GlnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPr 150  
1527 CGAGCTAAAGTAACACAAAGCATTTGCTTTGTAACCGCGAAGCTTCTCC 1478  
150 oTyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleA 167  
1477 TTATGCAAAAGTCTGGGGTCTAGGAGATGTTTGTGTTCATTGGCCAGTTG 1428  
167 laLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrrLeu 183  
1427 CTCCTTGCTGCTGCTGCTACCGCTGTGATGTTGTAATGCCAGATATTTA 1378  
184 AsnGlySerSerAspLysAsnTyrrAlaLysAlaLeuTyrrThrAlaLysHI 200  
1377 AATGTGTACCTCGATAAGAAATTTATCAATGCAATTTATACAGAAAAACA 1328  
200 stleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG 217  
1327 CATTGGGATTCATGCTTTGGCGGTGAACATGAAGTTTACCTTCTTCCATG 1278  
217 luTyrrArgAspAsnValAspTrpValPheValAspHisProSerTyrrHis 233  
1277 AGTATAGAGATTACGTTGACTGGGTGTTTGTGTGATCATCCCTCATATCAC 1228  
234 ArgProGlySerLeuTyrrGlyAspAsnPheGlyAlaPheGlyAspAsnG1 250  
1227 AGACCTGGAATTTATATGGAAGATAAGTTTGGTCTTTGGTGAATAATCA 1178  
250 nPheArgTyrrThrLeuLeuCysTyrrAlaAlaCysGluAlaProLeuIleL 267  
1177 GTTCAGATTACACACTCTTTGCTATGCTGCATGTGAGGCTCCTTTGATCC 1128  
267 euGluLeuGlyGlyTyrrIleTyrrGlyGlnAsnCysMetPheValValAsn 283  
1127 TTGAATTGGGAGGATATATTATGGACAGAATTGCATGTTTGTGTCAAT 1078  
284 AspTrpHisAlaSerLeuValProValLeuLeuAlaLysTyrrArgPr 300  
1077 GATGGCATGCGAGTCTAGTGCCAGTCTCTTCTGTCGAAATATAGACC 1028  
300 oTyrrGlyValTyrrArgAspSerArgSerThrLeuValIleHisAsnLeuA 317  
1027 ATATGGTCTTTATAAGACTGCCGCAGCATTTCTTGTAAATACATATTTAG 978  
317 laHisGlnGlyValGluProIleSerThrTyrrProAspLeuGlyLeuPro 333  
977 CACATCAGGGTGTAGAGCCTGCAAGCACATATCTGACCTTGGGTTGCCA 928  
334 ProGluTrpTyrrGlyAlaLeuGluTrpValPheProGluTrpAlaArgAr 350  
927 CCTGAATGGTATGAGGCTCTGGAGTGGGTATTCCTGAATGGGCGAGGAG 878  
350 gHisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValV 367  
877 GCATGCCCTTCACAAGGTGAGGAGTAAATTTTTCAGAAAGGTGCAGTTG 828  
367 alThrAlaAspArgIleValThrValSerGlnGlyTyrrSerTrpGluVal 383  
827 TGACGACAGATCGAATCGTACTGTCTAGTAAGGGTTATTCTGTGGGAGTCT 778  
384 ThrThrAlaGluGlyGlyGlnGlyLeuAsnGlnLeuLeuSerSerArgLy 400  
777 ACAACTGCTGAAGGTGACAGGGCTCAATGAGCTCTTAAGCTCCAGAAA 728  
400 sSerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnP 417  
727 GAGTGTATTAACGGAATTTGTAATGGAATTGACATTAATGATTGGAACC 678  
417 roThrThrAspLysCysLeuProHisIleTyrrSerValAspAspLeuSer 433  
677 CTGCCACAGACAATGATCCCTGCTCATTTATTCTGTGATGACCTCTCT 628  
434 GlyLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVa 450

627 GGAAAGGCCAAATGTAAAGGTGCATTCGACAGAGGAGCTGGGTTTACCTAT 578  
450 lArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrrGlnL 467  
577 AAGGCTGATGTTCTCTGTGATGGCTTTTATTGGAAGGTTGGATTATCAGA 528  
467 ysGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAsp 483  
527 AAGCATTTGATCTCATCAACTTATCATACCAGATCTCATCGGGAAGAT 478  
484 ValGlnPheValMetLeuGlySerGlyAspProIlePheGlyGlyTrpMe 500  
477 GTTCAATTTGTCATGCTTGGATCTGGTGACCCAGAGCTTGAAGATTGGAT 428  
500 tArgSerThrGluSerSerTyrrLysAspLysPheArgGlyTrpValGlyP 517  
427 GAGATCTACAGAGTCGATCTTCAAGGATAAAATTTCTGGATGGTGGAT 378  
517 heSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeu 533  
377 TTAGTGTTCAGTTTCCACCAGATAACTGCCGGCTGCGATATATTGTTA 328  
534 MetProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrrAlaMetG1 550  
327 ATGCCATCCAGATTGCAACCTTGTGGTCTCAATCAGCTATATGCTATGCA 278  
550 nTyrrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrV 567  
277 GTATGGCACAGTTCTCTGTGTCTCATGCAACTGGGGGCTTAGAGATACCG 228  
567 alGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrp 583  
227 TGGAGAACTTCAACCCITTGGTGAGAAATGGAGAGCAGGGGTACAGGGTGG 178  
584 AlaPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAl 600  
177 GCATTGCGACCCCTTAACACACAGAAACATGTTGTGGACATTTGCGAACTGC 128  
600 aMetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysA 617  
127 AATATCTACATACAGGGAACACAACTCTCTCTGGGAAAGGCTAATGAAGC 78  
617 rgGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSe 633  
77 GAGGCATGTCAAAAGACTTTCACGTGGGACCATGCCGCTGAACAAATACGA 28  
633 rArgSerSerSerGlyProSer 640  
27 CAAATCTTCCAGTGGGCTTCA 6







OM of: US-09-674-824-2 to: Pending\_Patents\_NA\_New:\* out\_format : pfs  
Date: Mar 28, 2002 6:41 PM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:  
-MODEL=frameet\_p2n.model -DEV=xlp  
-O/cgn2\_1/USPTO.spool/US09674824/runat\_27032002\_094253\_18518/app\_query.fasta\_1.824  
-DB=Pending\_Patents\_NA\_New -QPMT=fastap -SUFFIX=rnnpn  
-GAPOP=12.000 -GAPEXT=4.000 -HINTMATCH=0.100 -LOOPCL=0.000  
-LOOPEXT=0.000 -CGAPOP=4.500 -CGAPEXT=0.050 -XGAPOP=10.000  
-XGAPEXT=0.500 -DELOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000  
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200  
-THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODEB=LOCAL  
-OUTFMT=pfs -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09674824.@CIGN1\_11618 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:  
Query: US-09-674-824-2  
Query length: 756  
Database: Pending\_Patents\_NA\_New:\*  
Database sequences: 4692538  
Database length: 194516981  
Search time (sec): 716.860000

Sequence Strd Orig zScore EScore Len Documentation  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-674-824-1 + 404.00 4999.57 1.1e-269 277  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-508-377-11 + 3930.00 4858.62 7.6e-262 2  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-952-677-11 + 3612.00 4485.99 5.6e-240 23  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-931-297-1 + 2809.50 3470.80 1.5e-184 24  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-508-377-13 + 2346.00 2883.55 7.8e-152 1  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-815-264-86353 - 1523.00 1859.65 8.4e-95  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-10-018-418-1 + 1203.00 1477.82 1.6e-73 293  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-10-044-543-25 + 1201.50 1477.66 1.6e-73 284  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-10-018-418-3 + 1200.50 1475.01 2.2e-73 284  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-325-448-1034 + 1198.00 1473.34 2.8e-73 284  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-09-952-677-5 + 1182.50 1452.76 3.9e-72 28  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-10-018-418-5 + 1154.50 1420.60 2.4e-70 210  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-10-044-543-5 + 1150.00 1414.08 5.5e-70 234  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-865-419A-14741 + 1068.50 1314.75 1.9e-64  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-873-402A-83353 + 1062.00 1316.05 1.6e-64  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-865-439A-20602 + 987.50 1216.40 5.7e-59  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-815-264-57700 + 891.50 1062.51 2.1e-50  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-09-961-077-25 + 887.50 1089.04 7.0e-52 2  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-10-044-543-13 + 875.50 1074.52 4.5e-51 21  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-09-595-328C-273 + 870.00 1067.32 1.1e-50 22  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-10-044-543-21 + 853.50 1047.70 1.4e-49 20  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-815-264-17951 + 848.00 1043.67 2.4e-49  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-819-091A-6343 - 844.00 1042.68 2.7e-49  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-865-439A-20314 + 840.00 1032.94 9.4e-49  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-10-044-543-1 + 839.50 1027.90 1.8e-48  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-873-402A-67388 + 833.50 1022.36 3.6e-48 220  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-815-264-93310 + 809.00 1003.47 4.1e-47  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-10-021-323-969 + 798.50 990.32 3.5e-46 5  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-325-448-4323 + 798.00 984.32 4.8e-46  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-10-018-418-37 + 788.00 953.69 2.4e-44 90  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-865-439A-118661 + 779.50 967.33 4.2e-4  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-974-300-757 + 761.00 936.27 2.3e-43 1  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-849-529A-17365 + 720.00 893.33 5.6e-41  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-985-678-83476 + 642.00 799.61 9.3e-36  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-118-990A-1 + 634.50 778.48 1.4e-34 1  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-985-678-88686 + 632.00 787.30 4.5e-35  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-696-664A-17004 + 620.00 768.56 5.0e-34  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-708-427-12951 + 591.50 719.12 2.8e-31

/cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq:US-08-895-611-1 + 589.00 662.25 4.2e-28  
/cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq:US-08-895-611D-1 + 589.00 662.25 4.2e-28  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-895-611D-1 + 589.00 662.25 4.2e-28  
/cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-10-044-543-9 + 568.00 689.42 1.3e-29  
/cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-865-439A-21857 + 552.00 688.04 1.5  
seq\_name: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:US-09-674-824-1  
seq\_documentation\_block:  
: Sequence 1, Application US/09674824  
: GENERAL INFORMATION:  
: APPLICANT: Lorz, et al  
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH CODE FOR ENZYMES DERIVED FROM WH  
: TITLE OF INVENTION: WHICH ARE INVOLVED IN THE SYNTHESIS OF STARCH  
: FILE REFERENCE: AGR 1998/M 205/ FLH514413-3848  
: CURRENT APPLICATION NUMBER: US/09/674,824  
: CURRENT FILING DATE: 2000-02-16  
: PRIOR APPLICATION NUMBER: PCT/EP99/03156  
: PRIOR FILING DATE: 1999-05-07  
: PRIOR APPLICATION NUMBER: DE 198 20607.0  
: PRIOR FILING DATE: 1998-05-08  
: NUMBER OF SEQ ID NOS: 6  
: SOFTWARE: PatentIn version 3.0  
: SEQ ID NO 1  
: LENGTH: 2771  
: TYPE: DNA  
: ORGANISM: Triticum aestivum  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: (280)...(2547)  
US-09-674-824-1

alignment\_scores:  
Quality: 4044.00 Length: 756  
Ratio: 5.349 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-674-824-2 x US-09-674-824-1 ..  
Align seg 1/1 to: US-09-674-824-1 from: 1 to: 2771  
1 MetAlaAlaThrGlyValGlyValAlaGlyCysLeuAlaProSerValArgLe 17  
280 ATGGCGCGGACGGCGTTCGGCGCGGGTGGCTCGCCCCAGCGTCGGCCT 329  
17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34  
330 CGCGCGCGATCCGGCGACGGCGCGCGCGTCCGCTCGTCGTCGCGCG 379  
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50  
380 CGCGGCTCCGGCGCTTGGCGCGGGGCGGTACGTCGCGCGAGCTCAGCAGG 429  
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnLeuAlaProProLe 67  
430 GAGGGCCCCGCGCGCGCGCGCGCGAGCAGCAGCACTGGCCCCGCGCT 479  
67 uAlaProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84  
480 CGTGCCAGGCTTCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 529  
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100  
530 CCCCAGCAGCGCGCGCTTCGGCGAGCGCGCGCGTGGGGAACTCGCGGCC 579  
101 AspLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117  
580 GACCTCTCTGCTGAAGGATTCGTAGGATTCATCGACAGCATATTCCT 629  
117 lAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProg 134  
|||||



630 GGCTCAAGTCAGCAGGATCTTGAGATCATGGATGCCAATGAGCAACCTC 679  
134 lnaLalysValThrSerIleValPheValThrGlyGluAlaAlaPro 150  
680 AAGCTAAAGTTACACGTAGCATCGTGTGTGTGACTGGTGAAGCTGCCT 729  
151 TyrAlalysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167  
730 TATGCAAAAGTCAGGGGGTGGAGATGTTGTGTTGCTTACCAATGTC 779  
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184  
780 TCTTGCTCTCGTGCATCCAGTCAGTGGTGTAAATGCCAAGATACTAA 829  
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200  
830 ATGGGTCTCTGATAAAACTATGCCAAGGCATATATACATGCCAAGCAC 879  
201 ileLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG1 217  
880 APTAAGATTCCATGCTTTGGGGATCATGAAGTGACCTTTTTCATGA 929  
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234  
930 GTATAGAGACAACGTCGATTTGGTGTGTCGATCATCCGTATATCACA 979  
234 rProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250  
980 GACCAAGGAAGTTATATGAGAGATAATTTGGTGCTTTGGTGATTAATCAG 1029  
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267  
1030 TTCAGATACACACTCTTTTGTCTATGCTGCATCGAGGCCCTACTAATCCT 1079  
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284  
1080 TGAATTTGGGAGGATATATTTATGGACAGAAATTCATGTTGTGTGAAGC 1129  
284 sPTPrHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300  
1130 ATTGGCATGCCAGCTTGTGCCAGTCTTCTTGTGTCGCAAAATATAGACCA 1179  
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317  
1180 TACGGTGTATACAGAGATTCGCGCAGCACCTTGTATATATAAATTTAGC 1229  
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334  
1230 ACATCAGGGTGTGGAGCCTGCAAGTACATATCTGATCTGGGATGCCCTC 1279  
334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350  
1280 CTGAATGGTATGGAGCTTTAGATGGTATTTCCAGAATGGGCAAGGAGG 1329  
351 HisAlaLeuAspLysGlyGluAlaValAlaAsnPheLeuLysGlyAlaValVa 367  
1330 CATGCCCTTGACAAAGGTGAGGAGTAACTTTTGAAGGAGCAGTGTGT 1379  
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValTr 384  
1380 GACAGCAGATCGGATGTGACCGCTCAGTCAGGTTTATTCATGGAGGTCA 1429  
384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400  
1430 CAACTGCTGAAGGTGGACAGGGCTCAATGAGCTCTTAAGCTCCCGAAAA 1479  
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417  
1480 AGTGATTTGAATGGAATGTAAATGGAAATTTGACATTAATGATTGGAAACC 1529  
417 oThrThrAspLysCysLeuProHisIstYrSerValAspAspLeuSerg 434  
1530 CACCACAGACAAGTGTCTCCCTCATCATTTATTTCTGTCGATGACCTCTCTG 1579

434 lYlYsAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450  
1580 GAAAGCCAAATCTAAAGCTGAATTCAGAAGAGTGGGTTTACCTGTA 1629  
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467  
1630 AGGAGGATGTTCTCTGATGGCTTTATTTGGAAGACTGGATTACCGAA 1679  
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484  
1680 AGCATTTGATCTCATTTAAATGCCATTCAGAGCTCATGAGGAGGACG 1729  
484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500  
1730 TGCAATTTGTCATGCTTGGATCTGGGATCCAAATTTTGAAGCTGGATG 1779  
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517  
1780 AGATCTTACCGAGTCGAGTTACAAGGATATAATCCGTGGATGGGTGGATT 1829  
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534  
1830 TAGTGTCCAGTTTCCACAGAATAACTGCAGGTTGCGATATATTGTAA 1879  
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550  
1880 TGCCATCGAGATTTGAACCTTGGGTCTTAATCAGCTATATGCTATATGCAA 1929  
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567  
1930 TATGGTACAGTCTCTGTAGTTCATGGAACTGGGGCCCTCCGAGACACAGT 1979  
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584  
1980 CGAGACCTTCAACCTTTTGTGCAAAAGGAGAGAGGGTACAGGGTGGG 2029  
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600  
2030 CGTGTCTACCGCTAACCGTGGACAGAGTGTGTGGGCATTCGCGAACCCGG 2079  
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617  
2080 ATGTGACATTCAGGAGCACAAAGCCGCTCTGGAGGGGCTCATGAAGCG 2129  
617 gGlyMetThrLysAspHisThrTrpAspHisAlaProSerSerThrSerA 634  
2130 AGGCATGACGAAAGACCATAGCTGGGACCCTATGCCCGGAGCATCCTCTC 2179  
634 rGSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgGly 650  
2180 GATCTTCGAGTGGGCTTCTGGTGGACCAACCCTACGTCATGTAGAGGGGA 2229  
651 LeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSe 667  
2230 CTGGGAGGTCCTCAAGTGGCAGTCTCTTTCAGCTCTGAAGACATCCTCTC 2279  
667 rSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAlaT 684  
2280 ATCTCTCCCGCCCGCGAAGGATACCCCTGTACATTGCTGTCTCTGCTA 2329  
684 hrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThr 700  
2330 CAGTAGAGTCGCAATCGCCTCTCTTGTGTTTGGTTCGCGGTCGAGAAACA 2379  
701 TyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnle 717  
2380 TATGACGGCTGTCTGCTCGCGCGGTGACAGCTTCGGGTGGACGACAGTT 2429  
717 uGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAlaL 734  
2430 ACAGTTTGGGGAATAAGAAAGGATGTCTCGCAGGATGGTTAACAGCAA 2479



734 yshishisSerAspGlySerLeuSerValArgValThrAlaGluIleArg 750  
|||||  
2480 AGCACCACTCAGATGCGACCTCTCTGTCCGTGTACAGCTGAATCAGA 2529  
751 AsnGlnLeuValThrLeu 756  
|||||  
2530 AACCAACTGGTGACTCTT 2547

seq\_name: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq3:US-09-508-377-11

seq\_documentation\_block:  
: Sequence 11. Application US/09508377  
: GENERAL INFORMATION:  
: APPLICANT: KALEEN, ZHONGYILI  
: APPLICANT: MORELL, MATTHEW  
: APPLICANT: RAHMAN, SADBOUR  
: TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS  
: FILE REFERENCE: 054270/0126  
: CURRENT APPLICATION NUMBER: US/09/508, 377  
: PRIOR FILING DATE: 2000-06-09  
: PRIOR APPLICATION NUMBER: AU PP 2509  
: PRIOR FILING DATE: 1998-03-20  
: PRIOR APPLICATION NUMBER: PCT/AU98/00743  
: PRIOR FILING DATE: 1998-09-11  
: PRIOR APPLICATION NUMBER: AU PP 9108  
: NUMBER OF SEQ ID NOS: 71  
: SOFTWARE: PatentIn Ver. 2.1  
: SEQ ID NO 11  
: LENGTH: 2662  
: TYPE: DNA  
: ORGANISM: Triticum tauschii  
US-09-508-377-11

alignment\_scores:  
Quality: 3930.00 Length: 757  
Ratio: 5.247 Gaps: 1  
Percent Similarity: 98.943 Percent Identity: 98.415

alignment\_block:

US-09-674-824-2 x US-09-508-377-11 ..

Align seg 1/1 to: US-09-508-377-11 from: 1 to: 2662

1 MetaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17  
|||||  
257 ATGCGCGGACGCGCGTGGCGCGCGGTGCTCGCCCCACGCTCGCGCT 306  
17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34  
|||||  
307 CGCGCGCGATCCGCGGACGCGCGCGCGGTGCTCGCCCCACGCTCGCG 356  
34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50  
|||||  
357 CGCGCGTCCGCGCGCTTGGCGCGCGCGCGGTACCTTCCGAGCTCAGCAG 406  
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnGlnLeuAlaProProLe 67  
|||||  
407 GAGGCT 456  
67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84  
|||||  
457 CGTGCCAGGCTTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 506  
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100  
|||||  
507 CCGCC 556  
101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117  
|||||  
557 GACCTCTCTCGAAGGATTCCTGAGGATTCATCGACGACATTAATGT 606

117 lalaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134  
|||||  
607 GGCTGCAAGTCAGCAGGATTCAGATCATGATGCGAATGCGAAGCACTC 656  
134 lnAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150  
|||||  
657 AAGTAAAGTTACAGTACGATCGTGTGTGTGACTGGTGAAGCTGCTCCT 706  
151 TyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAl 167  
|||||  
707 TATGCAAGTCAGGCGCGCTGGGAGATGTTGTGGTTCGTTACCAATGC 756  
167 aLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184  
|||||  
757 TCTTGTCTGCTGCTGCTCACCCTGTGATGGTTGTAATGCCAAGATACCTGA 806  
184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200  
|||||  
807 ATGGTCTCTGATAAAACTATGCAAAAGCATATACACTGGGAAGCAC 856  
201 IleLysIleProCysPheGlyGlySerHisGluValThrPhePheHisG 217  
|||||  
857 ATTAAGATTCCATGCTTTGGGGGATCACATGAAGTGACCTTTTTCATGA 906  
217 uTyrArgAspAsnValAspTrpValPheValAspHisProSerTyrHisA 234  
|||||  
907 GTATAGAGACACGTCGATTTGGGTGTTGTGCGATCATCCGTCATATCATA 956  
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250  
|||||  
957 GACCAGGAAGTTATATGAGGATAAATTTGGTCTTTTGGTGATAATCAG 1006  
251 PheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLe 267  
|||||  
1007 TTCAGATACACACTCTCTTGTGATGCTGCGAGGCGCCCACTAATCCT 1056  
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284  
|||||  
1057 TGAATTTGGGAGGATATATTATGGACAGATTGCTATGTTGTGTGAAGC 1106  
284 spTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgPro 300  
|||||  
1107 ATTGGCATGCCAGCTTGTGCCAGTCTCTTGTGCAAAATATAGACCA 1156  
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317  
|||||  
1157 TACGGTGTGTACAGATTCCCGCAGCACCTTGTATATACATAATTTAGC 1206  
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProP 334  
|||||  
1207 ACATCAGGGTCTGGAGCCTTGCAGTACATATCTGATCTGGGATTTGCCAC 1256  
334 roGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArg 350  
|||||  
1257 CTGAATGGTATGGAGCTTTAGAATGGGTATTTCAGAAATGGGCAAGGAGG 1306  
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValVa 367  
|||||  
1307 CATGCCCTTGACAAGGGTGAGCAGCTTAACCTTTTGAAGGAGGAGCTCGT 1356  
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384  
|||||  
1357 GACAGCAGATCGAATGTGACCGTCAGTCAGGTTATTTCATGGGAGGTCA 1406  
384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400  
|||||  
1407 CAACTGCTGAAGGTGGACAGGCGCTCAATGAGCTCTTAAGCTCCCGAAA 1456  
401 SerValLeuAsnGlyIleValIasnGlyIleAspIleAsnAspTrpAsnPr 417  
|||||  
1457 AGTGATTGAATGGAATTGTAAATGGAATTGCATTAATGATTCGAACCC 1506  
417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434



|||||  
1507 CACCACAGCAAGTGTCCCTCATATTCTGTCGATGACCTCTCTG 1556  
434 lYsAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450  
|||||  
1557 GAAGGCCAANTGAAGCTGAATTCAGAGGAGCTGGGTTTACCTGTA 1606  
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLy 467  
|||||  
1607 AGGAGGATCTCTCTCATGTGGCTTTATTGGAAGACTGGATTACCAGAA 1656  
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484  
|||||  
1657 AGGCAATTGATCTCATTAATAATGCCATTCAGAGCTCATCAGGAGGAGC 1706  
484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMet 500  
|||||  
1707 TGCAGTTTGTTCATGCTTGGATCTGGGCATCCAATTTTGAAGGCTGATG 1756  
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPh 517  
|||||  
1757 AGATCTACCGAGTCGATGTAACAGGATAAATCCGTGGATGGGTTGGATT 1806  
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534  
|||||  
1807 TAGTGTCCAGTTTCCACAGATAAATCGCAGGTTGCATATATTGTTAA 1856  
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550  
|||||  
1857 TGCCATCCAGTTTGAACCTTGTGCTTAATCATCAGCTATATGCTATGCAA 1906  
551 TyrGlyThrValProValValHisGlyThrGlyGlyLeuArgAspThrVa 567  
|||||  
1907 TATGTACAGTTCTCTGTAGTTTCATGGAACTGGGGCTCCGAGACACAGT 1956  
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTrpA 584  
|||||  
1957 CGAGACCTTCAACCTTTTGTGTGCAAAAGGAGAGGAGGCTACAGGGTGG 2006  
584 laPheSerProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAla 600  
2007 CGTCTCACCGCTAACCGTGCACAGATGTTGTGGCAATTCGGAACCCG 2056  
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617  
|||||  
2057 ATGTCGACATTCAGGAGCACAAGCGCTCTGGAGGGCTCATGAAGCG 2106  
617 gGlyMetThrLysAspHisThrTrpAspHisAla.ProSerSerThrSer 633  
|||||  
2107 AGGCATGACGAAAGACCATACGTGGGACCATGCCCGCGGAGTACGAGC 2156  
634 ArgSerSerSerGlyProSerTrpThrAsnProThrSerCysArgArgGl 650  
|||||  
2157 AGATCTTCGAATGGCGCTTCGTGGACCAACCTACGTCATGTAGACGGG 2206  
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667  
|||||  
2207 ACTGGGAGGTCGAAGCGCGGCTCTCTGAGCTCTGAAGACATGTTCT 2256  
667 erSerPheArgGlyProGluGlyTyrProCysThrLeuArgCysProAla 683  
|||||  
2257 CATCTCTCCGCGCGCGGAAGGATACCCCTGTACATTGGCTTGTCTGCT 2306  
684 ThrValGluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgTh 700  
|||||  
2307 ACAGTAGATGCGCAATGCCCTGCTTGCT.TGCTCCGCGGTTCCGAGCT 2355  
700 rTyrAspGlyCysAlaAlaAlaAlaValThrAlaSerGlyGlyArgGlnL 717  
2356 AGATGACCGCTGCTGCTGCGCGGCTGACAGCTTCGGGTGATGACAGT 2405  
717 euGlnPheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAla 733  
|||||

2406 TACAGTTTTGGGAATAAGGAGGATGTCTGCAGAGTGGTTAACAGCA 2455  
734 LysHisHisSerAspGlySerLeuSerValArgValThrAlaGluIleAr 750  
|||||  
2456 AAGCACCACTCAGATGGCAGCCTCTCTGCTCCGTGTACAGCTGAATCAG 2505  
750 gAsnGlnLeuValThrLeu 756  
|||||  
2506 AAACCAACTGGTGACTCTT 2524  
seq\_name: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq2:US-09-952-677-1  
seq\_documentation\_block:  
: Sequence 1, Application US/09952677  
: GENERAL INFORMATION:  
: APPLICANT: Block, Martina  
: Lorz, Horst  
: Lutticke, Stephanie  
: Walter, Lennart  
: Froberg, Claus  
: Kossmann, Jens  
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES  
: FROM WHEAT WHICH ARE INVOLVED IN STARCH  
: SYNTHESIS  
: NUMBER OF SEQUENCES: 9  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave  
: STREET: 1251 Avenue of the Americas  
: CITY: New York  
: STATE: New York  
: COUNTRY: United States of America  
: ZIP: 10020  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
: CURRENT APPLICATION DATA: US/09/952,677  
: FILING DATE: 14-Sep-2001  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 09/196,390  
: FILING DATE: 19-Nov-1998  
: APPLICATION NUMBER: DE 196 21 588.9  
: FILING DATE: 29-MAY-1996  
: APPLICATION NUMBER: DE 196 36 917.7  
: FILING DATE: 11-SEP-1996  
: APPLICATION NUMBER: PCT/EP97/02793  
: FILING DATE: 28-MAY-1997  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Haley, Jr., James F.  
: REGISTRATION NUMBER: 27,794  
: REFERENCE/DOCKET NUMBER: AGREVO-9  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (212) 596-9000  
: TELEFAX: (212) 596-9090  
: INFORMATION FOR SEQ ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 2239 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: cDNA to mRNA  
: HYPOTHETICAL: NO  
: ANTI-SENSE: NO  
: ORIGINAL SOURCE:  
: ORGANISM: Triticum aestivum L.  
: STRAIN: cv. Florida  
: HAPLOTYPE: ca. 21 d Caryopses  
: IMMEDIATE SOURCE:  
: LIBRARY: cDNA library in pBluescript sk (-)  
: CLONE: TaSSS  
: FEATURE:



```
; NAME/KEY: CDS
; LOCATION: 3..2017
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-952-677-1

alignment_scores:
  Quality: 3612.00      Length: 671
  Ratio: 5.383          Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-674-824-2 x US-09-952-677-1
Align seg 1/1 to: US-09-952-677-1 from: 1 to: 2239

      86 ThrGlnProLeuProAspAlaGlyValGlyGluLeuAlaProAspLe 102
      3 ACGGAGCGCCCTTGGCGAGCGCGGGTGGGGAATCGCGCCGACCT 52
102 uLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleValAlaA 119
53 CTGCTCGAAGGGATTGCTGAGGATTCATCGACAGCATATTGTGGCTG 102
119 laSerGluGlnAspSerGluIleMetAspAlaAsnGluInProGlnAla 135
103 CAAGTGACGAGGATCTGAGATCATGATGCGAATGAGCAACCTCAAGCT 152
136 LysValThrArgSerIleValPheValThrGlyGluAlaAlaProTyrAl 152
153 AAAGTTACAGCTAGCATCGTGTGTGTGACTGGTGAAGCTGCTCTTATGC 202
152 aLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIleAlaLeuA 169
203 AAAGTCAGGGGGTGGGAGATGTTGTGCTTGTGTTACCAATTGCTCTTG 252
169 laAlaArgGlyHisArgValMetValValMetProArgTyrLeuAsnGly 185
253 CTGCTCGTGTCCAGAGTGTGTTGTAATGCCAAGATACCTTAAATGGG 302
186 SerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLy 202
303 TCCTCTGATAAAACATATGCAAGGCATTATACACTCGGAAGCACATTAA 352
202 sileProCysPheGlyCysHisGluValThrPhePheHisGluTyrA 219
353 GATTCATGCTTGGGGGATGCATGAAGTGACCTTTTTCATGAGTATA 402
219 rGAspAsnValAspTyrPValPheValAspHisProSerTyrHisAtgPro 235
403 GAGACAACGTCGATTGGGTGTTGTGTCATCATCCGTCATATACAGACCA 452
236 GlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheAr 252
453 GGAAGTTTATATGAGATAATTTTGTGCTTTTGTGTAATAATCAGTTGAG 502
252 gTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluL 269
503 ATACACACTCCCTTGTGATGCTGTCATGCGAGGCGCCCACTAAATCCTTGAAT 552
269 euGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnAspTyr 285
553 TGGGAGGATATATTTATGACAGAAATTCGATGTTGTGTGACGATTGG 602
286 HisAlaSerLeuValProValLeuLeuAlaAlaLysTyrArgProTyrGl 302
603 CATGGCAGCCTTGTGCGAGCTCTCTTGTGCAAAATATAGACCATACGG 652
302 yValTyrArgSpSerArgSerThrLeuValIleHisAsnLeuAlaHisG 319
653 TGTTTACAGAGATTCCCGCAGCACCTTGTATTACATAATTTAGCACATC 702
319 InGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProProGlu 335
```

```
|||||
703 AGGCTGTGGAGCCTGCAAGTACATATCCTGATCTGGGATTGCCCTCCTGAA 752
336 TrpTyrGlyAlaLeuGluTyrPValPheProGluTyrPalaArgArgHisAl 352
|||||
753 TGGTATGAGCCTTTAGAAATGGGTATTTCAGAAATGGCAAGGAGGAGCATGC 802
352 aLeuAspLysGlyGlyAlaValAsnPheLeuLysGlyAlaValValThrA 369
803 CCTTGACAAGGGTGAGCAGCTTAACCTTTTCAAAGGAGCAGTTGTGACAG 852
369 laAspArgIleValThrValSerGlnGlyTyrSerTyrGluValThrThr 385
853 CAGATCGGATTGTGACCGTCAGTCAGGGTTATTTCATGGAGGTGCACAACT 902
386 AlaGluGlyGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLysSerVa 402
903 GCTGAAGGTGACAGCGGCTCAATGAGCTCTTAAGCTCCCGCAAAAGTGT 952
402 lLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTyrPAsnProThrT 419
953 ATTGAATGGAATTGTAATGGAATTGACATTAATGATTGGAACCCACCA 1002
419 hrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerGlyLys 435
1003 CAGACAAGTCTCCCTCATCATATTCTGTCGATGACCTCTCTGGAAG 1052
436 AlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGl 452
1053 GCCAAATGTAAGCTCAATTGCAAGAGGAGTTGGGTTTACCTGTGAAGGA 1102
452 uAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyI 469
1103 GGAATGCTCTCTGATTGGGCTTTATTGGAAGACTGGATTACCGAAGGCA 1152
469 leAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGln 485
1153 TTGATCTCATTAATGAGCAATCCAGAGCTCATGAGGAGGAGGACGTGCAA 1202
486 PheValMetLeuGlySerGlyAspProIlePheGluGlyTyrPMetArgSe 502
1203 TTTGTCATGCTTGGATCTGGGATCCAAATTTTGAAGGCTGGATGAGATC 1252
502 rThrGluSerSerTyrLysAspLysPheArgGlyTyrPValGlyPheSerV 519
1253 TACCGAGTCGAGTTACAAGATAAATTCCTGGATGGGTGGATTGATTAGTG 1302
519 alProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetPro 535
1303 TTCCAGTTTCCCAAGATAAATGAGGTTGGGATATATTGTTAATGCCA 1352
536 SerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGl 552
1353 TCGAGATTGAAACCTTGGCGTCTTAATCAGCTATATGCTATGCAATATGG 1402
552 yThrValProValValHisGlyThrGlyGlyLeuArgAspThrValGluT 569
1403 TACAGTTCTCTGTAGTTTCATGGAACCTGGGGCTCCGAGACACAGTCGAGA 1452
569 hrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTyrPalaPhe 585
1453 CTTTCAACCTTTTGGTGCAAAGAGAGAGGGGTACAGGGTGGCGGCTTC 1502
586 SerProLeuThrValAspLysMetLeuTyrPalaLeuArgThrAlaMetSe 602
1503 TCACCGCTAACCGTGGCAAGATGTTGGGCATTTGGCAATTCGGAACCGCATGC 1552
602 rThrPheArgGluHisLysProSerTyrPValGlyLeuMetLysArgGlyM 619
1553 GACATTCAGGAGGACACAAAGCCGCTCCTGGGAGGGGCTCATGAAGCGAGGCA 1602
619 etThrLysAspHisThrTyrPAspHisAlaProSerSerThrSerArgSer 635
|||||
```



```
1603 TGACGAAGACCATACGTGGGACCATGCCCGAGCAGTAGTACGAGCATCT 1652
636 SerSerGlyProSerTrpThrAsnProThrSerCysArgArgGlyLeuGl 552
1653 TCGAGTGGGCGCTTCGTGGACCAACCCCTAGCTCATGTAGAGGGGACTGGG 1702
652 YArgSerLysCysGluSerProSerAlaLeuLysThrSerSerSerSerP 669
1703 GAGGTCCAAGTGGGAGTCTCTTCAGCTCTGAAGACATCTCTTCATCTCT 1752
669 heArgGlyProGluGlyTyrProCysThrLeuArgCysProAlaThrVal 685
1753 TCCGCGCGCGGAGGATACCCCTGTACATTGGCTGTCTGTCTACAGTA 1802
686 GluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrYrAs 702
1803 GAGTCCGAATGCGCCTGCTTGTCTTGGTTCGCGCGGTTCGAGAAACATGA 1852
702 pGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGlnLeuGlnP 719
1853 CGGCTGTGCTGCTGCGCGGTGACAGCTTCGGGTGGACGACATTACAGT 1902
719 heTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAlaLysHis 735
1903 TTTGGGCAATAAGCAAGGATGCTGCTCAGGATGGTTAACAGCAAGCAC 1952
736 HisSerAspGlySerLeuSerValatqValThrAlaGluIleArgAsnGl 752
1953 CACTCAGATGCGACGCTCTGCTCCGTTTACAGCTGAAATCAGAAACCA 2002
752 nLeuValThrLeu 756
2003 ACTGGTGACTCTT 2015
```

seq\_name: /cn2\_6/ptodata/2/pna/us09\_NEW\_COMB.seq2:us-09-931-297-1

```
seq_documentation_block:
: Sequence 1, Application US/09931297
: GENERAL INFORMATION:
: APPLICANT: Jens Kossmann
: TITLE OF INVENTION: Nucleic acid molecules encoding soluble
: starch syntheses from maize
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: James F. Haley, Jr., c/o FISH & NEAVE
: STREET: 1251 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10020
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30 (EPA)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/931,297
: FILING DATE: 16-Aug-2001
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/192,909
: FILING DATE: <Unknown>
: APPLICATION NUMBER: DE 196 19 918.2
: FILING DATE: 17-MAY-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Haley, Jr., James F.
: REGISTRATION NUMBER: 27,794
: REFERENCE/DOCKET NUMBER: GFB-9
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-596-9000
: TELEFAX: 212-96-9090
: INFORMATION FOR SEQ ID NO: 1:
```

```
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2383 base pairs
: TYPE: nucleotide
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Zea mays
: TISSUE TYPE: endosperm
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 2..1950
: OTHER INFORMATION: /function= "starch synthesis"
: /product= "soluble starch synthase"
: SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-931-297-1
```

```
alignment_scores:
Quality: 2809.50 Length: 749
Ratio: 4.411 Gaps: 10
Percent Similarity: 85.047 Percent Identity: 73.565

alignment_block:
US-09-674-824-2 x US-09-931-297-1 ..
Align seg 1/1 to: US-09-931-297-1 from: 1 to: 2383
```

```
4 ThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeuAlaAs 20
: :::::|||||
41 TCGCGCGTGGCGCGCGTCCCTCCTC..... 67
20 pProAlaThrAlaAlaArgAlaSerAlaCysValValArgAla..... 34
: |||
68 CTCGCGCGCGCGCGCTGGCGCGCGCGCTCGCGCACCGCGCGCGCGC 116
35 ..ArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
: |||||
117 GCGCGCTCCACGCGCTGCTCGCGCGCGCGTGGCGGAGCTGAGCAGG 166
51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
: |||||
167 GAGGCGCGCGCGCGCGCGATGCCACCGCGCTGCTGGCGCGCGCGCT 216
67 uValProGlyPheLeuAlaProProProAlaProAlaGlnSerProA 84
: |||||
217 CGTGCGCGCGCTTCTCTCGCG...CGCGCGCGCGCGCGCGGTGAGCGCG 263
84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
: |||
264 CATTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGGGGGCTCTCGG 313
101 AspLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVa 117
: :|
314 GAA.....CCTGAAGGAGTTGCTGAAGTTCCATCATCAACACAGTAGT 357
117 lAlaAlaSerGlnAspSerGluIleMetAspAlaAsnGluGlnProG 134
: |||
358 TGTGGCAAGTGAGCAAGATTCTGAGATTGTGTTGGAAGAGGAGCAGCTC 407
134 lAlaLysValThrArgSerIleValPheValThrGlyGluAlaAlaPro 150
: :|
408 GAGCTAAGTAACACAAACATTGCTTTGTAAGTGGGAGAGCTCTCTCT 457
151 TyrAlaLysSerGlyLeuGlyAspValCysGlySerLeuProIleAl 167
: |||||
458 TATGCAAAAGTCTGGGGTCTAGGAGATGTTTGTGTTTCATTGCCAGTTGC 507
167 aleuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLeuA 184
: |||||
508 TCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 557
```



184 snGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysHis 200  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
558 ATGGTACCCTCCGATAGAAATTATGCAATGCAATTTACACAGAAAACAC 607  
201 IleLysIleProCysPheGlySerHisGluValThrPhePheHisG1 217  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
608 ATTCGGATTCCCATGCTTTGGCGGTGAACATGAAGTTACCTCTTCCATGA 657  
217 uTyrArgAspAsnValAspTyrValPheValAspHisProSerTyrHisA 234  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
658 GTATAGAGATTCCAGTGGCTGTTGTTGATCATCCCTCATATACACA 707  
234 rgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGln 250  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
708 GACCTGGAAATTTATATGAGATAAGTTTGGTGCTTTTGGTGATNAATCAG 757  
251 PheArgTyrThrLeuLeuCystyrAlaAlaCysGluAlaProLeuIleLe 267  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
758 TTCAGATACACACTCCTTTGCTATGCTGCATGTGAGGCTCCTTTGGTCT 807  
267 uGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValValAsnA 284  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
808 TCAATTGGGAGATATATTTATGGACAGAATTGCATGTTTGTGCAATG 857  
284 spTrpHisAlaSerLeuValProValLeuLeuAlaLysTyrArgPro 300  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
858 ATTGCGATGCCAGTCTAGTCCAGTCTCTTCTGCTGCAAAATATAGACCA 907  
301 TyrGlyValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAl 317  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
908 TATGTTGTTTAAAGACTCCCGCAGCATTTCTGTAATACATAATTTAGC 957  
317 aHisGlnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProp 334  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
958 ACATCAGGCTAGAGCTCGCAACACATATCCTGACCTTGGGTGGCAC 1007  
334 roGluTyrTyrGlyAlaLeuGluTyrValPheProGluTyrAlaArgArg 350  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1008 CTGAATGATATGAGACTCGAGTGGGTATTCCTGAAATGGCGAGGAGG 1057  
351 HisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValva 367  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1058 CATGCCCTTGCAAGGGTGAGCGAGTTAAATTTTGAAGGTGCAGTTGT 1107  
367 lThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValT 384  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1108 GACAGCAGATCGAATCTGACTGTCAAGGGTTATTCTATGGAGGTCA 1157  
384 hrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerSerArgLys 400  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1158 CAACCTGCTGAAGGTGGACAGGGCTCAATAGCTCTTAAGCTCCAGAAAG 1207  
401 SerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnPr 417  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1208 AGTGATTAAACGAATGTAAATGGAATGACATTAATGATTGGAACCC 1257  
417 oThrThrAspLysCysLeuProHisHisTyrSerValAspAspLeuSerG 434  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1258 TGGCAGACAGAAATGTATCCCTGTCTATTATTCTGTGATGACCTCTCTG 1307  
434 lLysAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProVal 450  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1308 GAAAGGCCAANTGTAAAGGTGCATTCGAAAGGAGCTGGGTATACCTATA 1357  
451 ArgGluAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLys 467  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1358 AGGCCTGATGTTCTCTGATGCTTTTATGGAAGATTGGATTATCAGAA 1407  
467 sGlyIleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspV 484  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1408 AGGCATTGATCTCAATCACTTATATCATACCAGATCTCATCGCGGAAGATG 1457

484 alGlnPheValMetLeuGlySerGlyAspProIlePheGluGlyTyrMet 500  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1458 TTCAATTGTCTATGCTTGGATCTGGTGACCCAGAGCTTGAAGATTGGATG 1507  
501 ArgSerThrGluSerSerTyrLysAspLysPheArgGlyTyrValGlyPh 517  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1508 AGATCTACAGAGTCGATCTTCAAGGATAAATTCGTGGATGGTGGATT 1557  
517 eSerValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuM 534  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1558 TAGTGTTCACAGTTTCCACCACTGCTGCTCAATCAGCTATATGCTATGCAG 1607  
534 etProSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGln 550  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1608 TGCCATCCAGATTGCACTTGTGCTCAATCAGCTATATGCTATGTTAA 1607  
551 TyrGlyThrValProValValHisGlyThrGlyLeuArgAspThrVa 567  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1658 TATGGCAGACTTCTGTGTCATGCACTGGGGGCTTACAGATACCGT 1707  
567 lGluThrPheAsnProPheGlyAlaLysGlyGluGluGlyThrGlyTyrA 584  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1708 GGAGAAGCTTCAACCTTTTCGGTGAGAAATGGAGACGAGGTACAGGTGGG 1757  
584 laPheSerProLeuThrValAspLysMetLeuTyrAlaLeuArgThrAla 600  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1758 CATTCGACCCCTTAACACAGAAACATGTTGTGCATTCGCAACTGCA 1807  
601 MetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMetLysAr 617  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1808 ATATCTACATACAGGGAACACAGCTCTCTCTGGGAGGGCTAATGAAGCG 1857  
617 gGlyMetThrLysAspHisThrTrpAspHisAlaPro.SerSerThrSer 633  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1858 AGGCATGTCAAAAGACTTACGCTGGGACCATCCGCTGAAACAATACGAAC 1907  
634 ArgSerSerSerGlyProSerTyrTrpAsnProThrSerCysArgArgG1 650  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1908 AAATCTTCCAGTGGGCTTCTATCGATCGACCCCTATGTCATGTAAAAAGG 1957  
650 yLeuGlyArgSerLysCys.GluSerProSerAlaLeuLysThrSerSer 666  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1958 A.....CCAAAGTGGTGGTCTCTTGAAGATCATCAGTTCATCATCC 1998  
667 SerSerPheArgGlyProGluGlyTyrPro.....CysThrLeuArgCy 681  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1999 TATAGTAAGCTGAATGATGAAGAAAAACCCCTGTACATTACATGAAGGC 2048  
681 sProAlaThrValGluSerGlnCysAlaCysLeuLeuTyrPheAlaGlyS 698  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
2049 AGACCGGCTATTGGCTCCATTGCTCCAATGCTCTGCTTTGGCTGCCT... 2094  
698 erArgThrTyrAspGlyCysAlaAlaAlaValThrAlaSerGlyGly 714  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
2095 .....TGCCCTCGATGGACCGGATCGAGTGAAGATCCA 2127  
715 ArgGlnLeuGlnPheTrpGlyIleArgLysGlyCysAla.....Al 728  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
2128 GCGAAGCAGCAGTTTGAAGGATAGGAAGGGAGCTGGGAAGCAGTCACGC 2177  
728 aGlyTrpLeuThrAlaLysHisHisSerAspGlySerLeuSer 742  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
2178 AGGCAGCCTCGCGTGATTTCATTCATGAACAAGCTGGAGTCAGT 2220

seq\_name: /cgn2\_5/ptodata/2/pna/US09\_NEW\_COMB.seq3:US-09-508-377-13

seq\_documentation\_block:

; Sequence 13, Application US/09508377

; GENERAL INFORMATION:

; APPLICANT: KALEEN, ZHONGYILI

; APPLICANT: MORELL, MATTHEW

; APPLICANT: RAHMAN, SADEOUR

; TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS



```

; FILE REFERENCE: 054270/0126
; CURRENT APPLICATION NUMBER: US/09/508,377
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: AU PP 2509
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: PCT/AU98/00743
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: AU PP 9108
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 10337
; TYPE: DNA
; ORGANISM: Triticum tauschii
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (10232)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-09-508-377-13

alignment_scores:
  Quality: 2346.00      Length: 3317
  Ratio: 3.141         Gaps: 25
  Percent Similarity: 22.520  Percent Identity: 22.339

alignment_block:
US-09-674-824-2 x US-09-508-377-13 ..
Align seg 1/1 to: US-09-508-377-13 from: 1 to: 10337

1 MetAlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLe 17
|||||
1 ATGGCGGCGACGCGGCGTGGCGCGCGGCGTCCGCCCGCCCGCGCCT 50

17 uArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysValValArgA 34
|||||
51 GCGCGCGGATCCGCGGACGCGCGCGCGCGCGCGCGTCCGCTTGGCTCGCGCG 100

34 laArgLeuArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArg 50
|||||
101 CGCGGCTCCGCGCGCTTGGCGCGCGCGCGCGCGCTAGCTCGCGGAGCTCAGCAG 150

51 GluGlyProAlaAlaArgProAlaGlnGlnGlnLeuAlaProProLe 67
|||||
151 GAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 200

67 uValProGlyPheLeuAlaProProProProAlaProAlaGlnSerProA 84
|||||
201 CGTGCAGGCTTCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 250

84 laProThrGlnProProLeuProAspAlaGlyValGlyGluLeuAlaPro 100
|||||
251 CCCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300

101 AspLeuLeuLeu.Glu..... 105
|||||
301 GACCTCCTGCTCGAGGTAACAAACAAAGGCTGAATCCTCAGATCACTCGG 350

105 ..... 105
351 CGTCTTCGTTTTACCAAAATACGCTACTGCGAAAGTGGTGTATATGTGA 400

105 ..... 105
401 AGTTTCTGCGATTTCTTCTGACGGATGTTTCAGTCGATTTCAGTTGTATA 450

105 ..... 105
451 TATGTGATACGTTGTTGTTTCATCGATCGTACAGATTTTACCAGCACACTA 500

105 ..... 105
```

```

501 GATAGAAATCGAGACCGACGCGGCGAGATCAATAGATTTTCTAGACGTT 550
105 ..... 105
551 TTATTGGATCGTGAGATGATTGTTGGGGTGGCGTGTGATACGATAGCG 600
105 ..... 105
601 GTGCACCGCGGATGTATCGGGGCGATGTGCACGTGGTGGGTCTCAGCAGA 650
105 ..... 105
651 CATATCACTAGACTGGTATCGTAATTTACTAGTACTACTGGAAGAGGAC 700
105 ..... 105
701 TAAAGGCTAGGCCAAGTGCACGCGATGTTGGGAACGTTGTTAAATTGAT 750
105 ..... 105
751 GAGTTTGTCTTGTCTTGGGCTGGTATATTACCAAAAAATGGTGTAGT 800
105 ..... 105
801 CCCTGTACTTATTAAATGGGAAATCTTAACATGACACTGGGTTTATGAG 850
105 ..... 105
851 TCTCCAATTGTATTTCTCAGCACTCAACTGATTTTACTGATACTGTAGT 900
105 ..... 105
901 GGAATGACACGCTGAGCAGCCCCCTTCAAGGATGCAATGCTTCTTTCG 950
105 ..... 105
951 TTTTATATTACGAACACTAGAAGGAGCTTCCACCTTTGAGTACAGAAGTA 1000
105 ..... 105
1001 CTCCTCGTTCCAAATAGATGACTCAACTTTGTACTAATTTTGTACTA 1050
105 ..... 105
1051 TAGTTAGTACAAAGTTGAGTCATCTATTTTGAACGGAGGAGTAGTATC 1100
105 ..... 105
1101 GAAATTGAAGACCTTGTATTACTGTCTTGTTTTCAATGAAATGGGAG 1150
105 ..... 105
1151 GCCCATGCAATAGTACATGGCACCTGGGAGGCTGGGATCATGTGTGC 1200
105 ..... 105
1201 TTTGACAGACTACTAGACCCAGCTCACCCCTCTGTTAGATTACTTGTGG 1250
105 ..... 105
1251 TGCTACTTTGTGTTGCTGTGAGTATATCAGACATCCTGAATTTGGCAT 1300
105 ..... 105
1301 CTAGCTGAGACAGAATCAGGTTGCACCATCTTATTATTGCTAAACTG 1350
105 ..... 105
1351 TTGTCACGCAATTTATAAAGAATGTGATCTTCTGAGTATTAAATTAATCAT 1400
105 ..... 105
```



1401 GTCTGCTAATATCTGTCTCTGCTGTGGTTGACAAATATACCATATGA 1450  
106 ..... GlyIleAlaGluaspSerIleAspSerIle 115  
1451 ATATTTTCCATTTTCCACCAGGGATTGCTGAGGATTCCTCGACAGCAT 1500  
115 eileValAlaLaSerGluInAspSerGluIleMetAspAlaAsnGluG 132  
1501 AATCGTGGCTGCAAGTGAGCAGGATTCTGAGATCATGATGCGAATGAGC 1550  
132 InProGlnAlaLysValThrArgSerIleValPheValThrGlyGluAla 148  
1551 AACCTCAAGCTAAAGTTACAGTACGATCGTGTGTTGTGACTGGTGAAGCT 1600  
149 AlaProTyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuPr 165  
1601 GCTCCTTATGCAAGTCAAGGGGCTGGGAGATGTTGTGGTTCTGTACC 1650  
165 orleAlaLeuAlaLaArgGlyHisArgValMetValValMetProArgT 182  
1651 AATTGCTCTTGTCTGCTGCTGCTACCGTGTGATGTTGTTAATGCCAAGAT 1700  
182 TyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAla 198  
1701 ACTTGAATGGTCTCTGATAAACACTATGCAAGGCCATTTATACATGCG 1750  
199 LysHisIleLysIleProCysPheGlyGlySerHisGluValThrPhePh 215  
1751 AAGCACATTAAGATTCCATGCTTTGGGGATCACATGAAGTGACCTTTT 1800  
215 eHisGluTyrArgAspAsnValAspThrVal...Phe..... 226  
1801 TCATGAGTATAGACACAACGTCGATTGGGTGGGTACACAAATCACTTCTT 1850  
226 ..... 226  
1851 ATTCTCTGTGAATTGTAGCACTGTTATATCCCTGTTTACACTCTTTTA 1900  
226 ..... 226  
1901 GCCCTGCAAGACATATGTGATTTCCTATCTTTTGTATTATCCCTTGT 1950  
226 ..... 226  
1951 ACTCTTGCTCATGAGGTCAAAATATCATATATCCATGGAAGTCATGCAT 2000  
226 ..... 226  
2001 GTCCCTAGTATTTTGTGTGCTGCTGCTTTAACTTTTCAGGGATTAAATAG 2050  
226 ..... 226  
2051 TGGAAATTGTAACTAAAGTTTATTTTATTTGAAAAAATTTGAGTTGGC 2100  
226 ..... 226  
2101 TGAGCCACAGCAGCAGTCAGTGGCACCACCTGCTTGCACATGATTTTGCATT 2150  
226 ..... 226  
2151 TCTGTTGACCGAGCAGCTTCATGTGAATAAGGTGTAAATATCAATAAGTA 2200  
226 ..... 226  
2201 CCAATTTTATCTGCCAATTGCACITTAAGAGTATATACATTTATCTTGGC 2250  
226 ..... 226  
2251 CTCAATCATGGGAGTACTGTGCATTTCAGTGCACCATCATTTGTTCAAGGA 2300  
226 ..... 226  
2301 GAAATGTGGGTGCAAGGAAGACACTTTTGTCCCTTAATAAAGGAGCGC 2350

226 ..... 226  
2351 ACTCTGTTGTCAATATAGATAGAAAGCAACAACCTTATTTCAAGAGCTAA 2400  
226 ..... 226  
2401 CAATGGCAAAAGAAACCAAAAGAGCATGCTAAGGCGGTGACACCAAAAGG 2450  
226 ..... 226  
2451 TGAGGGGGGCTTGTGACTGACAGCACCCCAAACTATTGCCATTGTTTTA 2500  
226 ..... 226  
2501 CTAATGAAGATCATTTTAGAAGCTCTCAGGAACCTTCGAAAAACAGTGGCT 2550  
226 ..... 226  
2551 TTCGGTCCACAGATCGTCTGCTTAATATTTTGTCCAGTGATACTTTTTT 2600  
226 ..... 226  
2601 GCTCCTTACAAGAGTGCCTATGTTGACATATACATTTGTTAAGTTGTTTCA 2650  
226 ..... 226  
2651 AAGTTTACTTCTTATTCTAAACAGCAAGTCCCTAATGCTTGCAATTTATT 2700  
226 ..... 226  
2701 TGCTATTTATTTTATTTCTCATTTCAATCAACACATTTTGTTCAGGTGTT 2750  
227 ValAspHisProSerTyrHisArgProGlySerLeuTyrGlyAspAsn 243  
2751 TGTCTGATCATCGTCTATATCATAGACCAGGAAGTTTATATGGAGATAAT 2800  
243 heGlyAlaPheGlyAspAsnGln.Phe..... 251  
2801 TTGGTCTTTTGGTGATAATCAGGTACACTACACTATATACTAAGCTCCTAG 2850  
251 ..... 251  
2851 TTGACTAAGTCGTAAGTTGTTACCTCTCGCTGACCGGCTGCTCTATGTCG 2900  
252 ..... ArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeu 265  
2901 TGCAGTTTCAGATACACACTCTCTTGTCTATGCTGTCATGGAGGCCCACTA 2950  
266 IleLeuGluLeuGlyGlyTyrIleTyrGlyGlnAsnCysMetPheValVa 282  
2951 ATCCCTGAATTGGGAGGATATATTATATGACAGAATTCATGTTGTTGT 3000  
282 IasnAspTrpHisAlaSerLeuValPro..Val..... 292  
3001 GAACGATTGGCATGCGACGCTTGTGCCAGTGTACGTTGTTGTGGATCTG 3050  
292 ..... 292  
3051 AAAGTCCCAATCCTTTATTCATCTCTGCTTTGTCAGTGTGCCCATGTCTAC 3100  
292 ..... 292  
3101 ATTTCTTTTATGCTTTTTCATGCTGTTCTTATATATTCATATATATGCTTA 3150  
292 ..... 292  
3151 TGGAGTCTAAAAGTTACCGGAGGAATAACTCTTAAGGATTTTCCTCAATC 3200  
292 ..... 292  
3201 AATTATCTTTAGCTTTAGTTPAACATTTACTGTGGCAACATAAATGTGTTT 3250



292 ..... 292  
3251 TGAGATTACAAGTTTCAAGATTCACACTTCACTAGTTGCTAGCTAATCTG 3300  
292 ..... 292  
3301 ATGTTTTCCCGGAGAAAATGCCATAAGCTTTGTCTCTGTATGATTTGATA 3350  
292 ..... 292  
3351 GAAAAGAGTTTATGTACACTCCCAAGAGGGGACCCAAATTTACACAC 3400  
292 ..... 292  
3401 CACACCCCTGAGAACTAGGCGCTGCCGAGAGACGGATGCAAGCCCACT 3450  
292 ..... 292  
3451 GCCCCTGCCTTAGCTCAAAGCGGCGCTCAGCTTGATTTGTCAAGTAAG 3500  
292 ..... 292  
3501 CTACAGTGTAGATTCGGCAAGTCGATTCGTCGAAGATGACAGTGTG 3550  
292 ..... 292  
3551 CGCTGCTTCCAAATCCACCAAACTATGAGCATGATCACTGGAGAAGTACC 3600  
292 ..... 292  
3601 TTTTCTCGCGCTGAGGGGTGGACTGGTGTCTGCTGCCAGTTTTC 3650  
292 ..... 292  
3651 AGATAAATCTGAAAAATGCATGTTTTGTATGATTTTAGTATCTTGGGACCC 3700  
292 ..... 292  
3701 TGGGTACCACCTAAGCTTTCACACAGTAATTTGCAAGTTACACCTATAAAA 3750  
292 ..... 292  
3751 GTAAGGTCATGATATGCATGTGTTTGGGTAGATCATGCTGCATGCATT 3800  
292 ..... 292  
3801 TTAGGAATTAGGACATGCCAGAACCGTGAGGCTTATGGGCAATTTCAT 3850  
292 ..... 292  
3851 TTGTTCCATTATACGAGTCATGAATATGGTTACAGCATGTTTGGACGCTAC 3900  
292 ..... 292  
3901 TTGTTTGGGCAATTTACAGTGGTGAATTTAGCTGCTTGATGTTGGCTA 3950  
292 ..... 292  
3951 GCTGGCTTATTTGTACAAGTATCGATGTAGATGCATATTTCCCTTTTGT 4000  
292 ..... 292  
4001 TCTTGTGCTGTTTGGCATGTTGTATTTCCCTTTTCTGCGGAGTGTTC 4050  
292 ..... 292  
4051 ATGTTAAATTCGCTTTTCATTACATAATCAACTTTGTTGCTGACATCAGTC 4100  
293 .....  
4101 ATTTTATTCAGCCTTCTTGTGCTGCAAAATATAGACCATACGGTGTTCAC 4150  
305 rgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisGln..... 319

4151 GAGATTCGCCGACACCCCTTGTATACATAATTTAGCACATCAGGTTTGG 4200  
319 ..... 319  
4201 GTCATACACCTTTTCATATCCGTACATGGCTTTGTAAAGTCGGTTCCACAG 4250  
320 ..... Gly.ValGluProAla 324  
4251 TATCGTCATACACTGTATGTTATTTCAATGTTCATAGGTTGTGGAGCTGCA 4300  
325 SerThrTyrProAspLeuGlyLeuProGluTrpTyrGlyAlaLeuGln 341  
4301 AGTACATATCCTGATCTGGGATTCACCTGAATGGTATGGAGCTTTAGA 4350  
341 uTrpValPheProGluTrpAlaArgHisAlaLeuAspLysGlyGlu 358  
4351 ATGGGTATTTCCAGATGGCAAGGAGGCATGCCCTTTGACAAGGCTGAGG 4400  
358 laValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValThr 374  
4401 CAGTTAACTTTTGAAGGAGCAGTTGTGACACACAGATCGAATTTGTGACC 4450  
375 ValSerGln..... 377  
4451 GTCAGTCAGGTGAATACTCAATACTTCTTTTTCITTTTTCGGGATGTT 4500  
377 ..... 377  
4501 CTTTCAGTTCAATTCGCCCTGTCTTTTACCCTAATTAAGAAATGATTTAATCT 4550  
378 ..... GlyTyrSerTrpGluValThrAlaGluGlyGln 390  
4551 TTTGTTTCTAGGTTTATTCATGGGAGGTCACAACTGCTGAAGTGGACAG 4600  
391 GlyLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGly.Ile. 406  
4601 GGCTCAATGAGCTCTTAAGCTCCCGAAAAAGTGTATTGAATGGTAACATA 4650  
406 ..... 406  
4651 TATTTGAATCCACTTATCTTCTTGAACATATTTACAGAAATAGATGG 4700  
407 ..... Va 407  
4701 ATGGTTGCAAGATAAATTCAGTTTGCCTCTTTCGGTATGAAGAAATGCT 4750  
407 IAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuP 424  
4751 AAATGGAATTTGACATTAAATGATTGGAACCCACACACAGACAAGTGTCTCC 4800  
424 roHisHisTyrSerValAspAspLeuSerGlyLys..... 435  
4801 CTCATCATTTATCTGTCGATGACCTCTCTGGAAGGTTGTGTGATAGTAC 4850  
435 ..... 435  
4851 CCTATATAATAACATGATATATCTGATCTAGTACTTTCTTTTCTTTGCTA 4900  
435 ..... 435  
4901 GTTTGCTTCCATGATGTTCTCACAATACTAATCCTATGTTGGCATA 4950  
436 ..... AlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuPro 449  
4951 CTTGTCAGGCCAAATGTAAAGCTGAATTTGCAGAAGGAGCTGGGTTTACCT 5000  
450 ValArgGluAspValProLeu..... 456  
5001 GTAAGGAGGATGTTTCTCTGTTAGATACAAACCCCTTAAGATATATATT 5050  
456 ..... 456



5051 TTTTAAATCCCTAAAAAACTTGGCGATCTCTCATTTAGCTTGATTCAC 5100  
457 ..IleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIle 472  
|||||TGGCTTTATTGGAGACTGGATTACCAAGAGGCATTGATCTCAT 5150  
473 LysMetAlaIleProGluLeuMetArgGluAspValGlnPhe..... 486  
|||||TGGCTTTATTGGAGACTGGATTACCAAGAGGCATTGATCTCAT 5150  
5151 AAAATGGCCATTCCAGACTCATGAGGGAGGACGTGCAGTTTGAAGTTC 5200  
486 ..... 486  
5201 ATATTCCTTTTCTTGAGACTAGAGTATATAATCAAAACATGTAGTGTGGG 5250  
486 ..... 486  
5251 TGGTATAATACAGACATAAGTTCCAGCTATTGCTTCCATGAGAAATTTAA 5300  
486 ..... 486  
5301 TGCTATTTCAGTAATATGCTACTGCAAGTTTGAACAAAGTTGAAGCAA 5350  
486 ..... 486  
5351 TAAATATATGTGTAGCACTGACCATGCAGTGCCTACTATAGCTGAATGTC 5400  
486 ..... 486  
5401 CTGTAGTCTATGTGATCTAACACACTCAACACATGTTTCGCATACAAA 5450  
486 ..... 486  
5451 CACATGCGTGGCGCAACAAACATACTCTACAATAAAATTTGGCTTGATGA 5500  
486 ..... 486  
5501 ACTGCAGACATGCTCTTATCTCTCCATTCACACATTTCTTGTTCACATTC 5550  
486 ..... 486  
5551 GCTGAAGACTAAGAGAGGGGACCAGGCGTGATGTAGCCAACTAGATCC 5600  
486 ..... 486  
5601 AGTAAGGAAGCTAGCCGAGCCTAGGAGGATTCGCTTAGGTAGTGAACG 5650  
486 ..... 486  
5651 TAGGGTCTCTGACAGGAAGCTTCGGAGCTAGTCGATGCAGTGGTGAGG 5700  
486 ..... 486  
5701 AGAGGTGTTGATATCCTTTTGGCTTCCAGAAACCAATGTAGGGACAGAA 5750  
486 ..... 486  
5751 GGCGAAGGAGGTGGAGGATACCGGCTTCAAGCTGTGTTACATGGGACGGC 5800  
486 ..... 486  
5801 TGCAACAGAAATGGGTAGGCATCTTGATCAACAGAGCCCTTAAGTATG 5850  
486 ..... 486  
5851 GAGTGTAGACGTCAAGAGACGTGGGACCGGATTATCCTCGTCAAGCTG 5900  
486 ..... 486  
5901 GTAGTTGGGACTTACTTCTCAATGTTATCAGCGTGTATGCCCCGCAAGT 5950  
486 ..... 486  
5951 AGGCCACATGAGAACGCCAAGAGGGAGTTCTGGGAAGCCCTGGAAGACA 6000

486 ..... 486  
6001 TGGTTAGGAGTGTACCGATTGGCGAGAAAGCTCTTCATAGGAGGAGACCTC 6050  
486 ..... 486  
6051 AATGGCCACGTGGGTACATCTAACATAGGTTTGAAGGGGCACATGGGG 6100  
486 ..... 486  
6101 CTTTGGCTATGGCATCAAGAATCAAGAAAGAGATGCTTTACGCTTTGCTC 6150  
486 ..... 486  
6151 TAGCTACGACATGATTGTAGCTAACACCCCTCTTTAGAAAGAGAGATCA 6200  
486 ..... 486  
6201 CATCTGCTGACTTTTAGTAGTGCCCAACACTAGCCAGATCGATTTTCATCC 6250  
486 ..... 486  
6251 TCTCGAAGAGAGAGATAGGTGTGCGCGCTAGACTGCAAGGTGATACCT 6300  
486 ..... 486  
6301 TCGGATTTCGTCCAGCGGATAAGCGTGCCAAAGTCGCTAGAAAGT 6350  
486 ..... 486  
6351 GGTGGAAGCTCAAGGGGAGGTAGCTCAGGCGTTCAAGGAGAGGTTCATT 6400  
486 ..... 486  
6401 AGGGAGGGCCCTTGGGAGGAGGAGGATGCGGCAATGTGTGGATGAA 6450  
486 ..... 486  
6451 GATGGCGACTTGCAATTCGTAAGTGGCCCTCGGAGGAGTGTGGAGTGTCCA 6500  
486 ..... 486  
6501 GGGGATGGAGAACGCAAGATAGGATACCTGCTGGTGGTAATGATGTC 6550  
486 ..... 486  
6551 CAGAAGGCAATTAAGAGAAAGATTTGCTTTAGACGCTATACTTGA 6600  
486 ..... 486  
6601 TAGGAGTGCAGTCAACATAGAAAAAGTACAAGATGGCGAAGAGCCGCAA 6650  
486 ..... 486  
6651 AGCGAGCTGTCAAGTCAAGCAAGGGGTGCGCATATGAGGATCTCTACCAA 6700  
486 ..... 486  
6701 CGGTTAGCACGCAAGGAAGGCGAAAGGACATCTATAAGATGGCCAAGAT 6750  
486 ..... 486  
6751 CCGAGAGAGGAAGACCGAGGATATTGGCCAAGTCAAAATGCATCAAGGA 6800  
486 ..... 486  
6801 TGGAGCAGACCAACTCTTGGTGAAGGACGAGGAGATTAAAGCATAGATGC 6850  
486 ..... 486  
6851 GGGAGTACTTCGACAAGCTGTTCAATGGGAGGATGAGAGTCTTACCATT 6900







```
8701 TATCAGCTTGGATATATATTATAATGTTCAAAACATTTATGTCCTCTTTT 8750
529 .....CysAspIleLeuLeuMetProSerArgPheGluProCysGlyL 543
8751 GTGCAGTTGCGATATATGTTAATCCATCCAGGTTTGAACCTTGTC 8800
543 euAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValHisGly 559
8801 TTAATCAGCTATATGCTATGCAATATGTACAGTTCCTGCTAGTTCATGGA 8850
560 ThrGlyGlyLeu.Arg ..... 564
8851 ACTGGGGCCCTCCGAGTAAGCAACTGCCTTGAAAAATTTATGTTATCTTG 8900
565 .....AspThrVal 567
8901 GCTCCAACGCAAAATGTTCTAATTTGGCTCGTGTATTCACAGGACACAGTC 8950
568 GluThrPheAsnProPheGlyAlaLysGlyGluGlyThr.Gly.... 582
8951 GAGACCTTCAACCCCTTTTGGTGCAAAAGGAGAGGAGGTACAGGTACGC 9000
582 ..... 582
9001 ACTGCTCAATTTAGCTAACTTTTCAGTTTATCTTTTTCGAATGCTTGGG 9050
582 ..... 582
9051 GGTTCATTGCCCAATAAATCACTTGTGATAATTAACCTGTTACTGTTCTG 9100
583 .....TrpAlaPheSerProLeuThrValAspLysMetLeuTrp. 595
9101 TACTTGAGGTGGCGGTTCTCACCGCTAACCGTGGACAAGATGTTGTGGG 9150
595 ..... 595
9151 TAAGTTTTTGTGAGCTCTTGTCCGGTTATAGGATCGACCTTGCTGTAG 9200
595 ..... 595
9201 CATGGTACCTTAGTGCCCTTGTATATAGACCTAACCTGATGGACTCACT 9250
595 ..... 595
9251 TTGCTACACTAATCATAGTAGTCGATTGCCGGAGCGGTTTGTGTGGA 9300
595 ..... 595
9301 TTCTGCTAATTTAATTTTCATGACGATACTCATACATGGTTTGGTTCT 9350
595 ..... 595
9351 CCGATGGGGCCAGAAATGGCGTCTAGTGTCTGCGATCTGTGTAACTAGCC 9400
596 .....Ala. 596
9401 AATGCCGGGTGTTCCAAAGTGAATTTACCTTTTGACCAATTTGTCAGGC 9450
597 LeuArgThrAlaMetSerThrPheArgGluHisLysProSerTrpGluG 613
9451 ATTGCGAACCGCATGTCGACATTCAGGGAGCACAAGCCGCTCTGGGAGG 9500
613 lLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHisAla.Pr 629
9501 GGCTCATGAACGAGGCGATGACGAAAGACCATACGTGGGACCATGCCGC 9550
629 oSerSerThrArgSerSerGlyProSerTrpThrAsnProThrS 646
9551 GAGCAGTACGACGATCTTCGAATGGGCTTCGTGGAGCAACACCTTACCT 9600
646 erCysArgArgGlyLeuGlyArgSerLysCysGluSerProSerAlaLeu 662
9601 CATGTAGACGGGACTGGGGAGGTGCAAGCGGGGTCTCTCTTGAGCTCTG 9650
```

```
663 LysThrSerSerSerPheArgGlyProGluGlyTyrProCysThrLe 679
9651 AAGACATGTTCTCTATCTTCCGCGCCCGAAGATACCCCTGTACATT 9700
679 uArgCysProAlaThrValGluSerGlnCysAlaCysLeuLeuTrpPheA 696
9701 CGGTGTGCTCTGTACAGTAGAGTCGAATGCGCTGCTTGTCT.TGGTCCG 9749
696 laGlySerArgThrTyrAspGlyCysAlaAlaAlaValThrAlaSer 712
9750 CCGGTCGAGAGTAGATCAGGCTGTGCTGCTGCGCGGTGACAGCTTCG 9799
713 GlyGlyArgGlnLeuGlnPheTrpGlyLeArgLysGlyCysAlaAlaG 729
9800 GGTGATGACAGATTACAGTTTGGGGAATAAGGAAGGATGTGCTGCAGG 9849
729 yTrpLeuThrAlaLysHisSerAspGlySerLeuSerValArgValT 746
9850 ATGGTTAACAGCAAGCACCACTCAGATGGCAGCCTCTCTGTCGCTGTTA 9899
746 hrAlaGluIleArgAsnGlnLeuValThrLeu 756
9900 CAGCTGAATCAGAAACCAACTGGTGACTCTT 9931
```

seq\_name: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq1:US-09-815-264-86353

```
seq_documentation_block:
; Sequence 86353, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 86353
; LENGTH: 16095
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(16095)
; OTHER INFORMATION: unsure at all n locations
; US-09-815-264-86353
```

alignment\_scores:  
Quality: 1523.00 Length: 1887  
Ratio: 3.058 Gaps: 27  
Percent Similarity: 26.391 Percent Identity: 23.953

alignment\_block:

US-09-674-824-2 x US-09-815-264-86353/rev ..

Align seg 1/1 to reverse of: US-09-815-264-86353 from: 1 to: 16095

```
159 AspValCysGlySerLeuProfileAlaLeuAlaArgGlyHisArgVa 175.
16093 GATGTTTGGTTTCACTGCCAATTCCTTCTGCTCTGTTGGTCATCGTGT 16044
175 lMetValValMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrA 192
```



|||||GATGGTGTAAATCGCGAGATACATGAACGGGCTTGACCAAAATTTTG 15994  
192 laLysAlaLeuTyrThrAlaLysHisIleLysIleProCysPheGlyGly 208  
15993 CAACGCATTTTACACTGAGAGACACATTAAGATTCCATGCTTTGGCGGA 15944  
209 SerHisGluValThrPhePheHisGluTyrArgAspAsnValAspTrp.. 224  
15943 GAACATGAAGTTACTTTTTTTCACGAGTATAGGATTTCTGTTGATGGGT 15894  
224 ..... 224  
15893 GAGTATTCATCTCCTTACTATTTTCTGATGAACAAACATATCTGCTAT 15844  
224 ..... 224  
15843 CTTTTTGTCTTATCCCACTGAAGACATGATCATATATCTGTAAT 15794  
224 ..... 224  
15793 CTGTTTCTATGCTTACTGCTATACACTTGTGCACCTGGTTGATGCTGG 15744  
224 ..... 224  
15743 TGTTCAGATCAAAATGCAAAAGATTCCCTTCATATGACTTAAATGTATCC 15694  
224 ..... 224  
15693 APTTTATGCTCAAGCATTTGTGACTTCTGTTTGCACAAGCACTTAATA 15644  
224 ..... 224  
15643 TTAATAAGTTGTACTGTTGGCTTTCCATTCATTAATAGTCACAGTA 15594  
224 ..... 224  
15593 ACATTTAGTTATATCATGGAATTTATGTAATCTCAATGCAATTTTATTT 15544  
224 ..... 224  
15543 CTTGAAAGTGAACATGTGGTTACATTGAAACATGTATGTGCACAACATGG 15494  
224 ..... 224  
15493 TAATAGATGTAGATGCGATGACATGATATCATTAAGGAACAAGGATGTG 15444  
224 ..... 224  
15443 TATAACCTAGGAGGAACATTTCTAGCATTTTGGTTGCAAGATTGATTC 15394  
224 ..... 224  
15393 CTTGTTGAGTTTACCAATGTTTACTTATTCCTATTTGGAATTTGGTAGTG 15344  
224 ..... 224  
15343 TTCACCTCCGCGACCATCACCCATGACACATGCTGATTGTAGAAAG 15294  
224 ..... 224  
15293 CATACCTGCAATTTGGGTACTTATTTTGTGTACACCTCCTTCCTTG 15244  
225 .....ValPheValAspHisProSerTyrHisArgProGlySerLeuTyr 239  
15243 TTCAGGTGTTGTTGATCATCCCTCATATCATAGACCTGGAATTTGTAT 15194  
240 GlyAspAsnPheGlyAlaPheGlyAspAsn..... 249  
15193 GGAGATAAATTTTGGTGTCTTTTGGCGATAATTCAGGTATCTTACATCAAGA 15144  
249 ..... 249

15143 TCCTATTTGACCAGTCTAAATTTATATACATTCAGTTCATTGATTCGACTT 15094  
250 .....GlnPheArgTyrThrLeuLeuCystYrAlaAla 260  
15093 GACTTCCCTATATACAGTGCAGTATACACACTCTCTGTGCTATCGGGG 15044  
261 CysGluAlaProLeuIleLeuGluLeuGlyTyrIleTyrGlyGlnAs 277  
15043 TGTGAAGCCCATTAATCTTGAACCTGGGAGGATATATCTATGGACAGAA 14994  
277 nCysMetPheValValAsnAspTrpHisAlaSerLeuValProVal..Leu 293  
14993 ATGCATGTTTTGTTGTGAATGATGGCATGCCAGCTTGTGCCAGTGTATG 14944  
293 ..... 293  
14943 TTCTGGATCTGAACCTCTGCAACCAAAATTGATTAAACATTGAATATCTCTT 14894  
293 ..... 293  
14893 ATAAACCATTTTTTTCATACGTTTTGCTTTGTGTTGTGTGTCTATGTT 14844  
293 ..... 293  
14843 AGATAAGAGGAACAAATATATTAGATTGAGTGAATAGGAGGATTTTCTCAA 14794  
293 ..... 293  
14793 TCATTTTTTTTCTCAAAAGCGCAGGAGAACTGGGATTCATTTTCATTATTT 14744  
293 ..... 293  
14743 TCTCAATCATTTTGATATTTCCAAAAACATGAACAGTCCAGTACATCATTT 14694  
293 ..... 293  
14693 TATGTGTTTAAATGATTTAGTTTTGTGCTACACTAGTTTCATAGCTGATCT 14644  
293 ..... 293  
14643 GAAATCTCCTGCCTATGTTTTTTTTTGGAAATATTTCTGAGCTAGTGC GA 14594  
293 ..... 293  
14593 AGTTCTTTATTTAGTGTAGTGCCTATTTTGTAGTTCTCTTACAGATACA 14544  
293 ..... 293  
14543 ATTCAATACCCAAAGGATTTGGGAAAAATCCTGTAGTCTAAATGGCCTAT 14494  
293 ..... 293  
14493 CATGCTACTAGGTGACTGAGGTGATAGTTTCCAACCCCTCTAATTTGTGGT 14444  
293 ..... 293  
14443 AAATTTAACTGCTGTTTTCACACCATAATTTGATGTTTATAATTATTAAG 14394  
293 ..... 293  
14393 AGAGTGGTGGTACATGCTTTCTAACTTCTAAGACACTTCAGAAACACAT 14344  
293 ..... 293  
14343 GAGGTTAATGGTGAATGCTTTCCCTGCATATATTAATATGTTTTCAT 14294  
293 ..... 293  
14293 CCATCAGATGCTACATTCGCTACTCCTCTTTTGGCATGTTCTCCACCGT 14244  
293 ..... 293  
14243 GGATAAGATACTAATGCGCAGTGAAAAACTCATCACACTTATTTTTTAGTC 14194



293 ..... 293  
14193 CCTTTTACTCAATGCTGAATGTTAAATTTGTTTTCCTTACATATCAAT 14144  
294 ..... LeuAlaAlaLysTyrA 299  
14143 TTAATATAACTTTTATCTGTTTATCCAGCCTTCTTCTGCAAAATATA 14094  
299 rGpTyrGlyValTyrArgAspSerArgSerThrLeuValIleHisasn 315  
14093 GACCATATGGTGTACAGGATGCCGCCAGTGTCTTGTCATACATAAT 14044  
316 LeuAlaHisGln..... 319  
14043 CTAGCACATCAGGTTTGAGCCTATCATTTTGCATATCTTTTAACTAGTTT 13994  
319 ..... 319  
13993 TGTAAATGCCACTTTGTAAAGTCCAGTATGTAGTATTTTCATTTGTGTTCAA 13944  
320 ..... GlyValGluProAlaSerThrTyrProAspLeuGlyLeuPro 333  
13943 TGTGATAGGTTGTGGAGCCTGCCAGTACATATCTGACCTGGGATTGCCA 13894  
334 ProGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgAr 350  
13893 CCTGAATGGTATGAGCATTAGAATGGTGTTCACAGTGGGCAAGCGC 13844  
350 qHisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGlyAlaValV 367  
13843 GCATGCCCTTGACAAGGGTGAGGCGATCAATTTTTTAAAGGCGCAGTTG 13794  
367 alThrAlaAspArgIleValThrValSerGln..... 377  
13793 TGACAGCAGATCGAATTTGTACTGTCAGCCAGGTGAGATTCCCTCTTTT 13744  
377 ..... 377  
13743 CGTCTACTAGTTGGCTTACGGCCACAGTCAAAATTTGAAATTTGAACCTA 13694  
377 ..... 377  
13693 ACCTTAGAGTTGATTTTTTTTCCCATCTAGTCTATTTTATGCTGG 13644  
377 ..... 377  
13643 CTTTAAACCGCAAATTTCTGTTCAATTTTGTCTTGCACTCAATTAAGGA 13594  
378 ..... Gly.TyrSerTrpGluValThrAlaG 387  
13593 ATCATTTATGTTTCATACCAGGGGTATTCATGGGAGGTCACAACTGCTG 13544  
387 luGlyGlyGlnGlyLeuAsnGluLeuSerSerArgLysSerValLeu 403  
13543 AAGGTGGCAAGCCCTCAATGAGCTCTTAAGTCCCGGAAGAGTGTATTG 13494  
404 Asn..... 404  
13493 AATGGTACCTATATTTGGGTTTTTATATTGTTTCTTGAGTCAATCTG 13444  
405 .....G 405  
13443 TAATAGGTGGATGATACAAATTAATCAATCTATTTCCTGGATGTAG 13394  
405 lyIleValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLys 421  
13393 GAATTCGAATGGAATTGACATTAATGATTGGAACCCATCCACAGACAAG 13344  
422 CysLeuProHisHisTyrSerValAspAspLeuSerGlyLys..... 435  
13343 TTTCCTCCCTTATCATTTATCTGTGTGATGACCTGTCCGGAAGGTGAAGTA 13294

435 ..... 435  
13293 TAGTACCATGTATATACAGATTGTCATCTCGTTTCATTTTTTCTTTGCA 13244  
435 ..... 435  
13243 CCTCGTCGTTTCTTTTACAGACTAATTTATCAACACCTCGAATTCATGTTT 13194  
436 Ala..LysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArg 451  
13193 AGGCCAAGTGTAAAGCTGAATTTGCAGAAGGAGCTGGGTTTACCTATAAGG 13144  
452 GluAspValProLeu..... 456  
13143 CCCGATGTGCTCTGCTGTAGATACAAATATTTATGAAATTTATTTATTTA 13094  
457 ..... IleGlyP 459  
13093 GTATTCTTTTATGGGCAGTTGGATTATTAGCTTTATTATACAGATTGGCT 13044  
459 heIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAla 475  
13043 TTATTTGGAGATTGGACTATCAAAAAGGCATTGATCTAATTAACACTGCC 12994  
476 IleProGluLeuMetArgGluAspValGlnPheVal..Met..... 488  
12993 ATTCCAGATCTCATGGGACAATATTCAATTCGTAAGTTCTGTTTCTTT 12944  
488 ..... 488  
12943 TTTCTCGTGAGACTAGAAAGTCAATGTGATAGATCTTAGTGTATATAGT 12894  
488 ..... 488  
12893 AGTATATTGACACTTATATAGCTTTGTTATCATGAATCTAAGGAATGTAAT 12844  
488 ..... 488  
12843 ATGCAGCTATTGCTTTTCATTGTGAACGCATATGCAAGTAAACTAGTATTTT 12794  
488 ..... 488  
12793 CAGATCAAAATATTAGTGTCTATAAGAAAAAATGCATATGCAAGTCATAA 12744  
488 ..... 488  
12743 AGAAGATAGCAGTAGTATAACTTGACCACCTGTATGACTTATTAGCAACC 12694  
488 ..... 488  
12693 ATCCAGTGTGATATCAATTTGGAGCACCTGCTTGGCCAGCCAGTCTGTTAT 12644  
488 ..... 488  
12643 TTCCACCTTTTATACATTTNTGGTACACTPAAACAACATCTTCTGACAGAC 12594  
488 ..... 488  
12593 ACCAGGGAGCGGTATACATGCAGACAAAAGGATCACCATTATTATTAAGTC 12544  
488 ..... 488  
12543 AGTAGGATCTAACTTTGAGGACGAAATGCTTTGTATGAATGCAACACAG 12494  
488 ..... 488  
12493 TGCCTTATCCCAATACTTCATTGTTGGGGGTTTCTGGCTTTTCAACTGGAG 12444  
488 ..... 488  
12443 ACTAAACACAGAGTAGTGTGTTGAGTGGGTTTAAAAAATTTATCCACATAT 12394  
488 ..... 488



12393 TTTAGCCACCACCTGAATCGATTGATCAATGACACCTTTTAACTACTTCT 12344  
488 ..... 488  
12343 TTTTCTTAAAAAATGACACCTTTTGTGTTTGTGATAGGTCATG 12294  
489 LeuGlySerGlyAspProIlePheGluGlyTropMetArgSerThrGluSe 505  
|||||.....  
12293 CTTGGATCTGGTACCAGGTTTGAAGGATGGATGAGATCCACAGATC 12244  
505 rSerTyrLysAspLysPheArgGlyTropValGlyPheSerValProValS 522  
|||||.....  
12243 AGGTCACAGGATAAATTCGTGGATGGTGGATTTAGTTCAGTTT 12194  
522 erHisArgIleThrAla. Gly. .... 528  
|||||.....  
12193 CCCCGAATAACTGCAGGCTATGCAAAATAACTTCGCCAATAATACTCA 12144  
528 ..... 528  
12143 TTTGTCAGCTAAGATGCAATTTACATTTGAATCTCAATTAATTTTGGTG 12094  
529 .....CysAsp 530  
|||||.....  
12093 ATAAAGTTACATTTGAAATATTGATGCCCTCTATTTTGTACAGTTGGCAT 12044  
531 IleLeuLeuMetProSerArgPheGluProCysGlyLeuAsnGlnLeuTy 547  
|||||.....  
12043 ATATTGTTGATGCCATCCAGATTCGAACCTTGTGGCCTCAATCAGCTATA 11994  
547 rAlaMetGlnTyrGlyThrValProValIleHisGlyThrGlyGlyLeuA 564  
|||||.....  
11993 TGCATGCAATATGTCAGTGCCTGTTGTTTCATGGAATCGAGGCCTCA 11944  
564 rg..... 564  
11943 GAGTAAGCTTTCCTATAGTGTGACACTTATGTCAGCAGCAGCCCAAGGC 11894  
565 .....AspThrValGluThrPh 570  
|||||.....  
11893 AGAGCTGCTCACTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 11844  
570 eAsnProPheGlyAlaLysGlyGluGlyThrGly..... 582  
|||||.....  
11843 TAAACCGTTTGTCTGAGAAAGGAGACAGGTACAGGTAAGGAATGCTTC 11794  
582 ..... 582  
11793 ATTAAGTAAATTCGACATTTTATTTATTTATTTATTTATTTATTTATTTATTT 11744  
582 ..... 582  
11743 CAATTTTGTATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 11694  
583 .....TropAlaPheSerProLeuThrValAspL 592  
|||||.....  
11693 GCTACTTGGACATATTGACAGTGGGATCTCTGCCACTAACCAATTGAAA 11644  
592 yMetLeuTrp..... 595  
|||||.....  
11643 AAATGCTGTGGTAAAGTTATGTTGAACTGCACCTAGTTTATGTTCTTC 11594  
595 ..... 595  
11593 TATCCATTCCTTAATCCCATTTATATATATGCCATACACATCACTACTATTG 11544  
595 ..... 595  
11543 GCCCTCTTAACATCAATAGCATGATTTCAACTATGCAATTTTGGTTGT 11494  
595 ..... 595

11493 ATCGTATACTGCTATTTTCTTCCCGGTAAATCAATATGATAGTTGATTTCT 11444  
595 ..... 595  
11443 TGCACAAGTAAAAAGGATAAAAACAGGTGAGGCTCATGACGTATGTTCC 11394  
595 ..... 595  
11393 AGCCAGTTCTGGTTATTTGCTGCTGATTTCCAGTAGGATTTGAACCTCCTAGTA 11344  
595 ..... 595  
11343 CTTGTTAGTTTTCAGATTTAAGTAAATATTTAAGAAAGAGTTGACAAGC 11294  
595 ..... 595  
11293 AAATAAAGAAAGAACACACACGATTTGAATCCAAAGAAAGAAAGCGGC 11244  
595 ..... 595  
11243 TCAGCACTTGGTGGTTAATGGGAACCTGGAATTTAGTGGGAATAGAAAC 11194  
595 ..... 595  
11193 CTGAAGCATATATGCTTAGGTTGCTTAAATTTGCTCGACTGACCAGTTCCA 11144  
595 ..... 595  
11143 TGATTTCAATTTGCATGTTTGGTACATATCTCGAATGTGGAACACAGGAAA 11094  
595 ..... 595  
11093 CAGGCCCCATTTTGTGTCATTTGAACACTTCAATGAAACTGATATAAAAA 11044  
595 ..... 595  
11043 TGGATTTGACAAAGAAAGATTTGCAATTTGCTTGTCTACAATTTAATGTGAG 10994  
595 ..... 595  
10993 TACTAGTCAGTTTGTGCTTGTACTCAAGTAGTCAAGATAAATTT 10944  
595 ..... 595  
10943 AGGAAAGCTACAAACTTTGTTGAGAAACATTTAGAACAGTAATAAAA 10894  
595 ..... 595  
10893 AGCCTCATTTGAAATGGTATTTTCAGGSCACAGCAATTACACCATATGAA 10844  
596 .....Ala...LeuArg 598  
|||||.....  
10843 CTTTGGAGTATGACTAAAAAATTTGCAATCTGTTTTCAGGCATTCGGG 10794  
599 ThrAlaMetSerThrPheArgGluHisLysProSerTrpGluGlyLeuMe 615  
|||||.....  
10793 ATGGCAATTTGCACATACAGGACACAAAGTCTCTTGGGAGGCTCTAAT 10744  
615 LysArgGlyMetThrLysAspHisThrTrpAspHisAla.ProSerSer 631  
|||||.....  
10743 GAACGAGGCATGTCAACGCACTTTTACATGGACCATGCGCCCTCACAGT 10694  
632 ThrSerArgSerSerGlyProSerTrpThrAsnProThrSerCysAr 648  
|||||.....  
10693 ATGAACAGATCTTCGAATGGGCTTTCATGGATCAACCATATGTCATGTAA 10644  
648 g.ArgGlyLeuGlyArgSerLys...CysGluSerPro..... 659  
|||||.....  
10643 ATGGATTTGAAGGAGAGCAGCAATTTTCCGAGGAGCCCTCAATCTTCTG 10594  
660 .....SerAlaLeuLysThrSerSerSerPhe.ArgGly.....671  
|||||.....  
10593 TCTTTCATGCGGGAATCAAACTTTGTACACTACATGAAAGGGAACCA 10544



```

672 .....ProGluGlyTyrProCysThrLeuArg 680
10543 GTTATGCAAGTTGCAACGATCACTCAAGTTACCTTGTAGG..... 10500
681 CysProAlaThrValGluSerGlnCysAlaCysLeuLeuTrpPhe 695
10499 ...CCTGCTACTGGCCA.....ATATGGTTC 10476
seq_name: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:US-10-018-418-1

```

```

seq_documentation_block:
: Sequence 1, Application US/10018418
: GENERAL INFORMATION:
: APPLICANT: Morell, Matthew
: APPLICANT: Li, Zhongyi
: APPLICANT: Rahman, Sadequr
: APPLICANT: Appels, Rudolph
: TITLE OF INVENTION: Genes Encoding Wheat Starch Synthases and Uses Therefor
: FILE REFERENCE: 127-01
: CURRENT APPLICATION NUMBER: US/10/018,418
: CURRENT FILING DATE: 2001-10-29
: PRIOR APPLICATION NUMBER: PCT AU00/00385
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: AU P00052/99
: PRIOR FILING DATE: 1999-04-29
: NUMBER OF SEQ ID NOS: 54
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 2939
: TYPE: DNA
: ORGANISM: Triticum aestivum
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (176)..(2569)
US-10-018-418-1

```

```

alignment_scores:
Quality: 1203.00      Length: 815
Ratio: 2.522          Gaps: 22
Percent Similarity: 58.528      Percent Identity: 37.301

```

```
alignment_block:
```

```
US-09-674-824-2 x US-10-018-418-1 ..
```

```
Align seg 1/1 to: US-10-018-418-1 from: 1 to: 2939
```

```

8 AlaGlyCysLeuAlaProSerValArgLeuArgAlaAspProAlaThrAl 24
|||||
606 GCGGGTCCCGCCACCGCCGCGCGC.....CCCAGCCGTCG 643
24 aAlaArgAlaSer...AlaCysValValArgAlaArgLeuArgLeuA 40
|||||
644 AGCCAGAACAGATACCGGTTGAATGGTGAACAAAGCTAACGTCGCCTC 693
40 LaArgGlyArgTyrValAlaGluLeuSerArgGluGlyProAlaAlaArg 56
|||||
694 GCCCGCCAGGACG.ATAGCCGAGTCCGCGCTCCGGATCCCGCAGCTACC 742
57 ProAlaGlnGlnGlnGlnLeuAlaProProLeuValProGlyPheLeuAl 73
|||||
743 ATTTCCATCAGTGAACAGGCCAGAGTCGCTGCTCCACCGGAGAGGC 792
73 aProProProAlaProAlaGlnSerProAlaProThrGlnPro.... 88
|||||
793 GCCCGCGTCCGCGGCTCAAAATTTGCGCCCTCGGCTTCGCTCCCGGGT 842
88 .....
843 CTGACACTGTGACGCGCTGGAACCTTGAACCTGAAGAGGTCGCGTCATT 892
89 .....ProLeuProAspAlaGlyValGlyGluLeuAlaProAs 101

```

```

893 GTCAAAGAAGTCCAAACCCAAAGGCTCTTTTCGCGCCGCCGAGCACCC... 940
101 pLeuLeuLeuGluGlyIleAlaGluAspSerIleAspSerIleIleVala 118
941 .....GCTGTACACACAGACCTTTGGGACTTCGAAGAATACA 977
118 laAlaSerGluGlnAspSerGluIleMetAsp..... 128
978 TTGGTTTCGAGGAGCCCGTGAGGCCAAGGATGATGGCCGGGCTGTTGCA 1027
129 .....AlaAsnGluGlnProGln..... 134
1028 GATGATCGGGGCTCTTCGGAACACACACCAATCAGCATTCGGGCGCTTT 1077
135 ...AlaLysValThrArgSerIleValPheValThrGlyGluAlaAlap 150
1078 GGCAGGGGAGAACCTCATGAACGTGGTCTGCTGCTGCTGAATGTTCTC 1127
150 roTyrAlaLysSerGlyGlyLeuGlyAspValCysGlySerLeuProIle 166
1128 CCTGGTCAAAACAGGTGGTCTTGAGATGTTGCCGGTGTTCGCCAAG 1177
167 AlaLeuAlaAlaArgGlyHisArgValMetValValMetProArgTyrLe 183
1178 GCTTTGGCGAAGAGGACATCGTGTATGTTGTGTGTACCAAGGTAT.. 1225
183 uAsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrAlaLysH 200
1226 .....GGGACTATGAGGAAGCTACGATGTCGGAGTCGGAAT 1265
200 isIleLysIleProCysPheGlyGlySerHisGluValThrPhePheHis 216
1266 ACTACAAG.....GCTGCTGGACAGGATATGGAAGTGAATATTTCAT 1309
217 GluTyrArgAspAsnValAspTrpValPheValAspHisProSerTyr.. 232
1310 GCTTATATCATGAGTGTGATTTTGTTCATTGACGCTCTCTCTTCG 1359
233 .HisArgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheGlyAspA 249
1360 ACACCCCGCAGGAAGACATTTATGGG.....GGCAGCAGACAGAAA 1400
249 snGlnPheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAlaProLeu 265
1401 TTATGAAGCCATGATTTGTTCTGAAGCCGCTGTCGAGGTTCATGG 1450
266 IleLeuGluLeuGlyTyrIleTyrGlyGln...AsnCysMetPheVa 281
1451 CAGGTTCCATGCGCGGTGTCCTTATGGGATGGAAATCTGCTGTTAT 1500
281 lValAsnAspTrpHisAlaSerLeuValProValLeuLeuAlaAlaLysT 298
1501 TGCAATGATGGCAGCACGCACTCTGCTGCTATCTGAAAGCATAT 1550
298 yrArgProTyrGlyValTyrArgAspSerArgSerThrLeuValIleHis 314
1551 ACAGGACCATGTTTGTGTCAGTACACTCGTCCATTATGGTGATACAT 1600
315 AsnLeuAlaHisGlnGlyValGluProAlaSerThrTyrProAspLeuG 331
1601 ACATCGCTCACCAGGCGGTGGCCAGTAGATGAGTTCCTCCGCTTCCCGA 1650
331 yLeuProProGluTrpTyrGlyAlaLeuGluTrpValPheProGluTrpA 348
1651 GTTGCCTGAGCAGCTACTG.....GAACACT 1676
348 laArgArgHisAlaLeuAspLysGlyGluAlaValAsnPheLeuLysGly 364
1677 TCAGACTGTACGACCCCGTGGTGGTGAACACGCCCAACTACTTCGCCGCC 1726
365 AlaValValThrAlaAspArgIleValThrValSerGlnGlyTyrSerTr 381

```



```
1727 GGCCTGAAGATGCGGACAGGTTGTGCTGAGCCCGGGTACCTGTG 1776
381 pGluValThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuSerS 398
1777 GGAGCTGAAGACGCTGGAGCGGCTGGGGCTTCACGACATCATACGGC 1826
398 eArgLysSerValLeuAsnGlyIleValAsnGlyIleAspIleAsnAsp 414
1827 AGAAGCACTGGAAGACCGCGGCATCGTGAACGGCATCGACCAACATGGAG 1876
415 TrpAsnProThrThrAspLysCysLeuProHis.....HisTy 427
1877 TGAACCCCGAGGTGGACGTCACCTCAAGTCGACGGCTACACCAACTT 1926
427 rSerValAspLeu...SerGlyLysAlaLysCysLysAlaGluLeuG 443
1927 CTCCCTGGGGAGCTGGAGCTCGGCAACGGCAGTGCAGGAGGCCCTGC 1976
443 InLysGluLeuGlyLeuProValArgGluAspValProLeuIleGlyPhe 459
1977 AGCGGAGCTGGCCCTGCAGGTCCGCGCGACGTGCTCGCTGCTCGGCTTC 2026
460 IleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIle 476
2027 ATCGGGCGCTGGACGGCGCAAGAGGGCTGGAGATCATCGCGGACGGCAT 2076
476 eProGluLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyA 493
2077 GCCCTGGATCGTGACCGAGGACGTCGCGGCACTTCGAGCGGGAGCACCGAC 2126
493 spProfilePheGluGlyTrpMetArgSerThrGluSerSerTyrLysAsp 509
2127 GCCACGACCTGGAGGCGATGCTGCGGCACTTCGAGCGGGAGCACCGAC 2176
510 LysPheArgGlyTrpValGlyPheSerValProValSerHisArgIleTh 526
2177 AAGGTGCGGGGTGGGTGGGTTCCTCGCGGTGGCGCGACCGGATCAC 2226
526 rAlaGlyCysAspIleLeuMetProSerArgPheGluProCysGlyL 543
2227 GCGCGCGCGCGAGCGGCTCTCATGCCCTCCGGTTCGAGCGGTGCGGAC 2276
543 euAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValHisGly 559
2277 TGAACCACTCTACGCCATGGCTACGCGCACCGTCCCGCTCGTCATGCC 2326
560 ThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheGlyAlaL 576
2327 GTCGGTGGCGCTGAGGACACCGCTGCGCGGCTTCGACCCCTTC..... 2368
576 sGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrValAspLysM 593
2369 .AACCACTCGGCTCGGTGGAGTTCGACCGCGCAGAGCGCAGAGC 2417
593 etLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisLysPro 609
2418 TGATCGAGGCGCTCGGGCACTGCTCCGACCTACCGGAGCTACAAGGAG 2467
610 SerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisThrTrpAs 626
2468 AGCTGAGGGGGCTCCAGAGCGCGCATGTCCAGGACTTCAGCTGGGA 2517
626 pHisAlaIleProSerSerThrSerArgSerSerGlyProSerTrpThr 642
2518 GCATGCGGCAAGCTCTACGAGGACGCTCTCGTCAAGGCCAAGTACCACT 2567
643 Asn.....ProThrSerCysArgArgG 650
2568 GGTGAACGCTAGCTGATGCCGCTCCAGCCCGCGCATGCGTCATGACAGG 2617
650 yLeuGlyArgSerLysCysGluSerProSerAlaLeuLysThrSerSerS 667
2618 ATCGAAT.....TGGCGATTGCGCACGACGAGGAG 2646
```

```
667 erSerPhe.....ArgGlyProGluGly..... 674
2647 GTCCATGGAGCGCGGCATCCCGAAGTACAGTGACATGAGGTGTGTGT 2696
674 ..... 674
2697 GGTGAGACGCTGATTCCTGATCTGGTCCGTAGCAGAGTAGCGGAGGTA 2746
675 .....TyrProc 677
2747 GGGNAGCGCTCCTTGTACAGGTATATGGGAATGTTTAACTTGGTATT 2796
677 ysThrLeuArgCysProAlaThrValGluSerGlnCysAlaCysLeuLeu 693
2797 GTAATTTGTTATGT.....TGTGTGCAT 2819
694 TrpPheAlaGlySerArgThr.....TyrAspGlyCysAlaAl 706
2820 TATTACAGAGGGCAACGATCTGCGCGCGCACCGGCCCACTGTTGGGC 2869
706 aAlaAlaValThrAlaSerGlyArgGlnLeuGlnPheTrp 720
2870 CGGTCGCACAGCAGCGCTGGATCCGACCGCTGGGCGCTGG 2912
seq_name: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:US-10-044-543-25
seq_documentation_block:
; Sequence 25, Application US/10044543
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: Novel Starch Synthese Polynucleotides
; TITLE OF INVENTION: and Their Use in the Production of New Starches
; FILE REFERENCE: 1144D
; CURRENT APPLICATION NUMBER: US/10/044,543
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/388,743
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 2418
; TYPE: DNA
; ORGANISM: Typha latifolia
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2418)
; US-10-044-543-25
```

```
alignment_scores:
  Quality: 1201.50      Length: 643
  Ratio: 2.814          Gaps: 15
  Percent Similarity: 66.407  Percent Identity: 42.457
alignment_block:
US-09-674-824-2 x US-10-044-543-25
Align seg 1/1 to: US-10-044-543-25 from: 1 to: 2418
42 GlyArgTyrValAlaGluLeuSerArgGluGlyProAlaAlaArgProAl 58
|||||:|||||:
325 GGGAAGTACAAAGC.....GGCGCGTTCCTAA 353
58 aGlnGlnGlnGlnLeuAlaPro.....ProLeuValP 69
|||:|||||:
354 TAATTATTCCCAATTGGCTCAAGATGATACATCTGAGAATCCACTGTGA 403
69 roGlyPheLeuAlaProProProPro..... 77
|||||:|||||:
404 ATTCCTTTGGAGGTTCTCCAAAAGATAACGTAGAACGTGTTCAATTTAG 453
```



```

78 ...AlaProAlaGlnSerProAlaProThrGlnProProLeuProAspAl 93
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
454 GTGAGACAATCAGCAGTAGATGCTTTTGAAGACCTGAAGAACCCAGTTT 503
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
93 aGlyValGlyGluLeuAlaProAspLeuLeuGluGly..... 106
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
504 GGGACGACGACGAGATCTCTCACCATTCTTGAAGCTGAATCGATG 553
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
107 .....IleAlaGluAspSerIleAspSer 114
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
554 GTCTAAAGCAGGAAGTCTGAAGATCTTGTAGAAGCAAAATGGATAGT 603
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
115 IleIleValAlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnG 131
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
604 GTACATGTC.....AAAGATGATTGAATCCTGGGAGAAATGA 644
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
131 uGlnPro.....GlnAlaLysValThrArgSerIleValPheValT 145
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
645 GGTTCCTCTCTTGGCTGGGGCAAAATGTCATGAACATCATAGTAGTTG 694
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
145 hrGlyGluAlaAlaProTyrAlaLysSerGlyGlyLeuGlyAspValCys 161
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
695 CTGCAGATGTCCTCTGTGTCACCAACAGGTGGCTTGGAGATGTTGCA 744
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
162 GlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValVa 178
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
745 GGAGCATTCGCCAAGGCTTTGGCCCAAGAGACATAGGCTCATGGTTGT 794
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
178 lMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAla 195
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
795 GGCACCAAGGTATGGAACATGCTGCAACCCCAAGATATAGGAGCCGCA 844
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
195 euTyrThrAlaLysHisLysIleProCysPheGlyGlySerHisGlu 211
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
845 AATACTACAAGGTTTCAT.....GGCAGCATATGGA 876
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
212 ValThrPheHisGlyTyrArgAspAsnValAspTyrPheValPheValAs 228
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
877 GTAACATATTCATGCTATATACGACGGTGTGGATTGTTTATGGA 926
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
228 pHisProSerTyr...HisArgProGlySerLeuTyrGlyAspAsnPheG 244
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
927 TAGTCACAGACTCCCTACCGGGGAATCGTATTATGAGGGAAC..... 972
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
244 lValAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAla 260
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
973 .....CGAGTGGATATCTTAAACGCTATGATTTGTTCTGCAAGCAGCT 1017
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
261 CysGluAlaProLeuIleLeuGluLeuGlyGlyTyrIleTyrGlyGln... 276
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1018 GTAGAGGTTCTTGGCATGTTCCATGTGCTGGCTTCTTTATGGAGATGG 1067
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
277 .AsnCysMetPheValValAsnAspTyrPheHisAlaSerLeuValProVal 293
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1068 TAAATTGGCTTTTCATACGAAATGATTGGCATACTGCTCTCTTGGCTGTT 1117
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
293 euLeuAlaAlaLysTyrArgProTyrGlyValTyrArgSerArgSer 309
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1118 ATCTGAAGGCATATATCGTGACAAATGGCTTGATCAAAATATGCTCGTCT 1167
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
310 ThrLeuValIleHisAsnLeuAlaHisGlnGlyValGluProAlaSerTh 326
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1168 GTTCTGGTAAATACACATACAGTACCCAGGCTGCTGGTCTGTAGATGA 1217
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
326 rTyrProAspLeuGlyLeuProProGluTyrTyrGlyAlaLeuGluTyrP 343
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1218 CTTCAAAATTTTGGCTTGGCGGATCATACTTGGACCTTTTCAGATTGT 1267
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
343 alPheProGluTyrAlaArgHisAlaLeuAspLysGlyGluAlaVal 359
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1268 ATGACCCCGTC.....GGAGGTGAACACCTC 1293
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
360 AsnPheLeuLysGlyAlaValAlaValThrAlaAspArgIleValThrVal 376

```

```

seq_name: /cyn2_6/ptodata/2/pna/US10_NEW_COMB.seq:US-10-018-418-3

```

```

1294 AATATTTTGTCTGGCTGGAAGACTGCTGACCGAGTGGTTACTGTTAG 1343
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
376 rGlnGlyTyrSerTyrPheGluValThrThrAlaGluGlyGlnGlyLeuA 393
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1344 CCAATGTTATGATGGAGCTGAAAACATCAGAAAGTGGTTGGGGCCTTAC 1393
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
393 snGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleValAsnGly 409
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1394 ATGAAATATAAATCAAGTAAGTCAAGGTTTCAAGGTATTGTTAAATGGC 1443
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
410 IleAspIleAsnAspTyrPheProThrThrAspLysCysLeuProHis... 425
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1444 ATTGATCAAAAGGAGTGGAGCCCAATTTGATGTGCACCTTAAATCGGA 1493
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
426 .....HistySerValAspLeuSer...GlyLysAlaLysC 438
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1494 TGGATACACAAATATTCTTAGATACCTTAGAGATGGTAAGCCAGTAT 1543
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
438 ysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAspVal 454
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1544 GTAAAGGCTGCTTTCAGCGAGAGTGGTCTGCCTGTTCTGTAATAATGA 1593
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
455 ProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLe 471
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1594 CCCATCATTCATTCATGGAGGTTAGACCACCAGAAAGCGTCGATCT 1643
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
471 uIleLysMetAlaIleProGluLeuMetArgGluAspValGlnPheValM 488
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1644 CATTCGCCAGGCCCATGCTTGGATTGTTCAGTCATGATGTTCAAGTAGTCA 1693
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
488 etLeuGlySerGlyAspProIlePheGluGlyTyrMetArgSerThrGlu 504
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1694 TGTTAGCCAGCGGAGGCAAGACCTTGAGAAATTTACTCAGGAACCTTGAG 1743
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
505 SerSerTyrLysAspLysPheArgGlyTyrPheValGlyPheSerProVa 521
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1744 GGTCAACACAGCGGACAAAGTTAGAGCATGGTTCATTTTCAGTAAAGAT 1793
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
521 lSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgP 538
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1794 GGCATATAGAAATACAGCAGTCCGACATCTCATGATGCTTCGAGGT 1843
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
538 heGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrVal 554
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1844 TTGAGCCATCGGGATTGAACAGCTTTTACGCAATGATGATGGAACCAT 1893
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
555 ProValValHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAs 571
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1894 CCAGTGTGATGCTGTTGGGGGCTTAGAGATACAGTCACTCAATTTGA 1943
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
571 nProPheGlyAlaLysGlyGluGluGlyThrGlyTyrAlaPheSerProL 588
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1944 TCCCTTC.....AACGAGTCTGCTTGGTTGGACCTTCGACAGG 1984
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
588 euThrValAspLysMetLeuTyrAlaLeuArgThrAlaMetSerThrPhe 604
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1985 CAGAGGCGAGAAAGCTGATCCATGCAATGAATACTGCTTGAATACATAC 2034
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
605 ArgGluHisLysProSerTyrPheGlyLeuMetLysArgGlyMetThrLy 621
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
2035 TGAATTTACAAGACAGTGGAAAGGGTCTTCAAAACAAGAGGGATGATGCA 2084
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
621 sAspHisThrTyrAspHisAlaPro.SerSerThrSerArgSerSerSer 637
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
2085 AGATCTTAGCTGGGTAATGCTGCTCAGCAATACGAGGATGTCCTTTG 2134
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
638 GlyProSerTyrThrAsnProThrSer 646
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
2135 CAGCCCAAGTACCATGGTGTCTTCT 2161
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

```



seq\_documentation\_block:  
: Sequence 3, Application US/10018418  
: GENERAL INFORMATION:  
: APPLICANT: Morell, Matthew  
: APPLICANT: Li, Zhongyi  
: APPLICANT: Rahman, Sadequ  
: APPLICANT: Appels, Rudolph  
: TITLE OF INVENTION: Genes Encoding wheat Starch Synthases and Uses Therefor  
: FILE REFERENCE: 127-01  
: CURRENT APPLICATION NUMBER: US/10/018,418  
: CURRENT FILING DATE: 2001-10-29  
: PRIOR APPLICATION NUMBER: PCT AU00/00385  
: PRIOR FILING DATE: 2000-04-28  
: PRIOR APPLICATION NUMBER: AU P00052/99  
: PRIOR FILING DATE: 1999-04-29  
: NUMBER OF SEQ ID NOS: 54  
: SOFTWARE: PatentIn ver. 2.0  
: SEQ ID NO 3  
: LENGTH: 2842  
: TYPE: DNA  
: ORGANISM: Triticum aestivum  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: (89)..(2485)  
: US-10-018-418-3

alignment\_scores:  
Quality: 1200.50 Length: 784  
Ratio: 2.686 Gaps: 20  
Percent Similarity: 57.015 Percent Identity: 37.628

alignment\_block:  
US-09-674-824-2 x US-10-018-418-3 ..  
Align seg 1/1 to: US-10-018-418-3 from: 1 to: 2842

```

7 GlyAlaGlyCysLeuAla.....ProSerValArgLeuArgAl 19
|||||
194 GGGCGCGGAGGCTGCACCTGGCGCGCGTGGCGCGGACGCGCAGCGCTCG 243
19 asppProAlaThrAlaAlaArgAla..... 27
|||||
244 CGACGAGGTGTGGCGCGCGCGCGCGCGCGGAGAGGACGCGAGGTGG 293
28 .....SerAlaCysValValArgAlaArgLeuArgAlaArg 41
|||||
294 ACGACGACGCGCGTCCGCGAGGACGCGCGCGCGCGCGCGCGGTGGCGGC 343
42 GlyArgTyrValAlaGlu.....LeuSerArgGl 51
|||
344 GCCACCAAGGTCCGCGGAGGAGGATCCCGTCAAGACGCTCGATCGCGA 393
51 u.....GlyProAlaAlaArgProAlaGlnGlnGlnLeuA 64
|||
394 CGCGCGGGAAGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 443
64 laProProLeuValProGlyPheLeuAlaProPro..... 76
|||
444 CCGGTCCA.....CCGAGTATGAACGGCAGCGCGGTGAACGTGAGAAC 487
77 .....ProAlaPr 79
|||||
488 AATCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
79 oAlaGlnSerProAlaPro...ThrGlnProProLeuProAspAlaGly 95
|||||
538 CGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 587
95 alGlyGluLeu.....AlaProAspLeuLeuGluGlyIle 107
|||||
588 AAACAAAGCTAACTGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCG 637

```

```

108 AlaGluAspSerIleAspSerIleIleValAlaAlaSerGluGlnAspSe 124
|||
638 GCTCCGGATTCCGCGCTACCATTTCCATCAGTGACAAGCGCGCGAGCT 687
124 rGluIle..... 126
|
688 CGTTGTCCAGCCGAGNAGCGCGCGCTCGCGCTCAAAATTCGTGG 737
127 .....MetAspAlaAsnGluGlnProGln 134
738 TCTCGGCTTCTGCTCCAGCGCTGGACATTCACAGCGATGTTGAACCTGA 787
135 AlaLysVal..... 137
788 CTGAAGAAGGGTGGGCTCATCTCGTGAAGAAGCTCCAAACCAAGGCTCT 837
137 ..... 137
838 TTCGCCGCTTGCAGCCCCCGCTGTACAAAGAGACCTTTGGGACTTCNAGA 887
137 ..... 137
888 AATACATTGGCTTCGAGGAGCGCGTGGAGGCAAGGATGATGGCTGGCT 937
137 ..... 137
938 GTTCAGATGATCGGGCTCTTTGAACATCACACAGAACCATGATTCGGG 987
138 .....ThrArgSerIleValPheValThrGlyGluA 148
988 ACCTTTGGCAGGCGAGAACGTCATCAACGTTGGTGGTCTGCTGCTGAAT 1037
148 laAlaProTyrAlaLysSerGlyLeuGlyAspValCysGlySerLeu 164
1038 GTTCTCCTGTCGCAAAACAGGTGCTTGGAGATGTTGGCGGTGCTTGG 1087
165 ProfileAlaLeuAlaAlaArgGlyHisArgValMetValValMetProAr 181
1088 CCCAAGGCTTTGGCGAGAGAGAGACATCGTGTATGTTGGTACCAG 1137
181 gTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAlaLeuTyrThrA 198
1138 GTAT.....GGGACTATGAGGAAGCTACGATGTCGGAGTCC 1175
198 laLysHisIleLysIleProCysPheGlyGlySerHisGluValThrPhe 214
1176 GAAATACTACAAG.....GCTGCTGGACAGGATATGGAAGTGAATAT 1219
215 PheHisGluTyrArgAspAsnValAspTrpValPheValAspHisProSe 231
1220 TTCCATGCTTATATCGATGGAGTTGATTTCTGTTTCATTGACGCTCTCT 1269
231 rTyr...HisArgProGlySerLeuTyrGlyAspAsnPheGlyAlaPheG 247
1270 CTTCCGACACCGCCAGGAAGACATTTATGGG.....GGCAGCAGAC 1310
247 lyAspAsnGlnPheArgTyrThrLeuLeuCysTyrAlaAlaCysGluAla 263
1311 AGGAATATTGAAGCGCATGATTTTGTCTGCAAGCGCGCTGTCAGGTT 1360
264 ProLeuIleLeuGluLeuGlyTyrIleTyrGlyGln...AsnCysMe 279
1361 CCTTGGCAGTTCCATCGCGCGGTGCTCCCTTATGGGATGGAATCTGCT 1410
279 tPheValValAlaAsnAspTrpHisAlaSerLeuValProValLeuAlaA 296
1411 GTTTATTGCAAAATGATGGCACACGCGCTCTGCTCTATCTGCAAG 1460
296 laLysTyrArgProTyrGlyValTyrArgAspSerArgSerThrLeuVal 312
1461 CATATTACAGGCGCCATGTTTGTATGTCAGTACACTCGGTCCATTATGGTG 1510
313 IleHisAsnLeuAlaHisGlnGlyValGluProAlaSerThrThrProAs 329

```



```

1511 ATACATAACATCGCGCACCAGGCGCGTGCCTAGATGAATTCCTCGTT 1560
329 pLeuGlyLeuProGluTrpTyrglyAlaLeuGluTrpValPheProG 346
1561 CACCGAGTTCCTGACCACTACTG.....G 1586
346 luTrpAlaArgArgHisAlaLeuAspLysGlyGluAlaValAsnPheLeu 362
1587 AACACTTCAGACTGTACGACCCCGTGGTGGTGAGCACCGCAACTACTTC 1636
363 LysGlyAlaValValThrAlaSpargIleValThrValSerGlnGlyTy 379
1637 GCGCGCGCTGAAGATGCGCGACCAAGTGTCTGTGTGAGCCCGGGTA 1686
379 rSerTrpGluValThrAlaGluGlyGlnGlyLeuAsnGluLeuL 396
1687 CCGTGGGAGCTCAACAGCGTGGAGGGCGCTGGGGCTTCAGGACATCA 1736
396 euSerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAspIle 412
1737 TAGCGCAGAACACTGGAAGACCGCGCATCGTCAACGGCATCGACAAC 1786
413 AsnAspTrpAsnProThrThrAspLysCysLeuProHis..... 425
1787 ATGGAGTGAACCCCGAGGTGGACGTCCACCTCAAGTCGGACGGCTACAC 1836
426 HistyrSerValAspAspLeu...SerGlyLysAlaLysCysLysAlaG 441
1837 CAACCTTCCTCGGGACGCTGGACTCGGCAAGCGGAGTGCAGAGGAG 1886
441 luLeuGlnLysGluLeuGlyLeuProValArgLysValProLeuIle 457
1887 CCCTCAGCGGAGTGGCGTGCAGGTCCGCGCGCGCTGCGGTGCTC 1936
458 GlyPheIleGlyArgLeuAspTyrglnLysGlyIleAspLeuIleLysMe 474
1937 GGTTCATCGCGCGCTCGAGCGGCGAGGCGGTGGAGATCATCGCGGA 1986
474 tAlaIleProGluLeuMetArgLysValGlnPheValMetLeuGly 491
1987 GGCATGCGCTGGATGCTGTGAGCGGAGGAGTGCAGTGTGATGCGGCA 2036
491 erGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSerTyr 507
2037 CCGGCGCGCCAGCACTGGAGAGCATGCTGCGGCACCTTCGAGCGGGAGC 2086
508 LysAspLysPheArgGlyTrpValGlyPheSerValProValSerHisAr 524
2087 CACGACAGGTGCGGGGTGGTGGGTCTCCGTGCGCTGGCGCACCG 2136
524 gIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluProC 541
2137 GATCAGCGCGGCGCGCAGCGCTCTCATGCCCTCCCGGTTCGAGCGGT 2186
541 ysGlyLeuAsnGlnLeuTyralaMetGlnTyrglyThrValProValVal 557
2187 CGGGTGAACAGCTTTACGCCATGCCCTACGCGACCGTCCCGCTCGT 2236
558 HisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheGl 574
2237 CACGCCGTGCGGGGTGAGGACACCGTGGCGCGCTTCGACCCCTTC.. 2284
574 yAlaLysGlyGluGlyThrGlyTrpAlaPheSerProLeuThrValA 591
2285 .....AACCACTCCGCGCTCGGGTGGAGTTCGACCGCGCGGAGCGC 2327
591 spLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHis 607
2328 ACAAGCTGATCGAGCGCTCGGCGACTGCCTCCGCACTACCTACCGGACTAC 2377
608 LysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisTh 624
||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

2378 RAGGAGACTGGAGGGGCTTCAGGAGCGCGCATGTGCGAGGACTTCAG 2427
624 rTrpAspHisAlaProSerThrSerArgSerSerSerGlyProSer 640
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2428 CTGGAGCATGCGGCAAGCTCTACGAGGAGCTCTCTCAAGGCCAAGT 2477
seq_name: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:US-60-325-448-1034
seq_documentation_block:
; Sequence 1034, Application US/60325448
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.R.
; TITLE OF INVENTION: Constitutive rice promoters
; FILE REFERENCE: 1360.026PRV
; CURRENT APPLICATION NUMBER: US/60/325,448
; CURRENT FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 4708
; SEQ ID NO 1034
; LENGTH: 2412
; TYPE: DNA
; ORGANISM: Oryza sativa
US-60-325-448-1034

alignment_scores:
Quality: 1198.00 Length: 709
Ratio: 2.680 Gaps: 24
Percent Similarity: 63.047 Percent Identity: 42.313

alignment_block:
US-09-674-824-2 x US-60-325-448-1034 ..
Align seg 1/1 to: US-60-325-448-1034 from: 1 to: 2412

2 AlaAlaThrGlyValGlyAlaGlyCysLeuAlaProSerValArgLeu.. 17
:|||||:|||||:|||||:|||||:|||||:|||||:
396 AGTCGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 445
18 .....ArgAlaAspProAlaThrAlaAlaArgAlaSerAlaCysV 31
|||||:|||||:|||||:|||||:|||||:|||||:
446 CAAGTAAGAACGGAGCGCTTCTCAGCGCGCGCGCGCGCGCGCGCTGC 495
31 alVal.....ArgAlaArgLeuArgLeuAlaArgGlyArg 43
:|||||:|||||:|||||:|||||:|||||:|||||:
496 TCAGGGAACGATCGGTCTACCGCGCGCGCGCGCGCGCGCGCTGCGC 542
44 TyrValAlaGluLeuSerArgGlyGlyProAlaAla.....Ar 56
:|||||:|||||:|||||:|||||:|||||:|||||:
543 CACGCGCGCGTGCACCATACGAAGCTCCCGCGCGCGCGCGCGCGCTGA 592
56 gProAlaGlnGlnGlnLeuAlaProPro...LeuValProGlyPhe 71
|||||:|||||:|||||:|||||:|||||:|||||:
593 TCCTTCCATCCGTAGACAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 642
72 LeuAlaProProProAlaProAlaGlnSerProAlaProThrGlnPr 88
|||||:|||||:|||||:|||||:|||||:|||||:
643 ACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCC 692
88 o...ProLeuProAsp.....AlaGlyV 95
||| ||| |||||
693 TCTCCCAAGCGCTGACAATTCGGAATTTGCGAGAGGATAAGAGCGCAAA 742
95 alGlyGluLeuAlaPro..... 100
||| ||| |||||
743 TTGTTGAGAGTGTCCGAAGCAAGCGCGCGCGCGCGCGCGCGCGCTTC 792
101 ..... 104
793 GCGGTAGACAGAGACGTGGGTTCCTCAAGAAATATTTGATCTGAACGA 842
104 uGluGlyIleAlaGluAspSerIleAspSerIleIleValAlaAlaSerG 121
|||||:|||||:|||||:|||||:|||||:|||||:

```



408 snGlylleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuPro 424  
||||| ||||| : : : : : ||||| : : : : :  
1697 ACGGCATCGACTACCGCGAGTGAACCCGAGGTGACGTGCACCTGCAG 1746  
425 His.....HsTyrSerValAspAspLeu...SerGlyLysAl 436  
:  
1747 TCCGACGGCTACGCCAAGTACACCGTGGCTCGCTGCATCTCCAGCNAAGC 1796  
436 aLysCysLYsAlaGluLeuInLYsGluLeuGlyLeuProValArgGluA 453  
: : : : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : : : :  
1797 CGGTGTCAAGCGCGCGCTGCAGCGAGCTGGGGCTGGAGTGCAGCGACG 1846  
453 spValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLYsGlyIle 469  
1847 ACGTGGCGCTCATCGGGTTCATCGGGCGGCTCGACGGCGAGAAGGTGTG 1896  
470 AspLeuIleLysMetAlaIleProGluLeuMetArgGluAspValGlnPh 486  
||||| ||||| ||||| ||||| : : : : : ||||| : : : : :  
1897 GACATATCGGCAGCGGATCCGTGGATCGCCGGCAGCAGCTGCAGCT 1946  
486 eValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArgSerT 503  
: : : : : ||||| ||||| ||||| : : : : : : : : : : : : : : :  
1947 GGTGCTGTGGCTCCGGCGCGCGACCTCGAGAGTGCATCTGCACCGGT 1996  
503 hrGluSerSerTyrLYsAspLysPheArgGlyTrpValGlyPheSerVal 519  
1997 TCGAGGCGCACACACAGCAAGTGC CGCGGTGGTGGGGTCTCGGTG 2046  
520 ProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSe 536  
||||| : : : : : ||||| ||||| ||||| ||||| ||||| ||||| : : : : :  
2047 AAGATGGCGCACCGGATCAGCGGGCGCGAGCTGTGCTATGATCGCGTC 2096  
536 rArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyT 553  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : : : :  
2097 CGGTGTGAGCGGTGCGGCTCAACCAAGCTCTACGCCCTGGCTACGGCA 2146  
553 hrValProValValHisGlyThrGlyLeuArgAspThrValGluThr 569  
2147 CCGTCCCGCTGTGCACGCGCTCGCGGGTGAAGGGACACCGTGTGGCG 2196  
570 PheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPheSe 586  
2197 TTCGACCGTTT.....GAGGACCGCGCTCGGTGGAGCTTCGTA 2237  
586 rProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSerT 603  
: : : : : ||||| ||||| : : : : : : : : : : : : : : :  
2238 CGCGCCCGAGCGCACAGCTCATCGAGGCGCTCGGCCACTGCTCGAGA 2287  
603 hrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMet 619  
2288 CGTACCGCAAGTACAAGGAGAGCTGGAGGGGTTCACAGTGC GCGCGGTG 2337  
620 ThrLysAspHisThrTrpAspHisAla.ProSerSerThrSerArgSerS 636  
2338 TCGCAGGACCTCAGCTGGGACCACCGCCGCGAGCTCTACGAGGAGGTCT 2387  
636 erSerGlyProSerTrpThrAsn 643  
||| ||||| ||||| |||||  
2388 TGCTCAAGGCCAAGT...ACCAAT 2407

```
seq_documentation_block:
; Sequence 5, Application US/09952677
```

GENERAL INFORMATION:  
APPLICANT: Block, Martina  
Lorz, Horst  
Luticke, Stephan  
Walter, Lennart  
Fronberg, Claus  
Kossmann, Jens  
TITLE OF INVENTION: NUCLEIC



FROM WHEAT WHICH ARE INVOLVED IN STARCH  
SYNTHESIS

NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10020

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/952,677  
FILING DATE: 14-Sep-2001

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/196,390  
FILING DATE: 19-Nov-1998  
APPLICATION NUMBER: DE 196 21 588.9  
FILING DATE: 29-May-1996  
APPLICATION NUMBER: DE 196 36 917.7  
FILING DATE: 11-Sep-1996  
APPLICATION NUMBER: PCT/EP97/02793  
FILING DATE: 28-May-1997

ATTORNEY/AGENT INFORMATION:  
NAME: Haley, Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: AGREVO-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090

SEQUENCE CHARACTERISTICS:

LENGTH: 2825 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Triticum aestivum L.  
STRAIN: cv. Florida  
TISSUE TYPE: ca. 21 d Caryopses  
IMMEDIATE SOURCE:  
LIBRARY: cDNA library in pBluescript sk (-)  
CLONE: pTAS1

FEATURE:

NAME/KEY: CDS  
LOCATION: 162..2559  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-952-677-5

alignment\_scores:  
Quality: 1182.50 Length: 825  
Ratio: 2.571 Gaps: 24  
Percent Similarity: 55.758 Percent Identity: 36.727

alignment\_block:

US-09-674-824-2 x US-09-952-677-5 ..

Align seg 1/1 to: US-09-952-677-5 from: 1 to: 2825

7 GlyAlaGlyCysLeuAla.....ProSerValArgLeuArgal 19  
|||||  
267 GGGCGCGGAGGTGACCTGGCGCGCGCGCGCGCGCGCGCGCTCG 316  
19 aAspProAlaThrAlaAla..... 25  
||| |||

317 CGACGGAGCTGTGCGCGCGCTCGCGCGCGCGGAGAGGACGCGGGGATCG 366  
26 ..ArgAlaSerAlaCysValValArgAlaArgLeuArgLeuAlaArg 41  
|||||  
367 ACAGCGCGCGCGCTCGGTGAGCAGCCCGCGCGCTCCCGCGGTGGCGCC 416  
42 GlyArgTyrValAlaGlu.....LeuSerArgG1 51  
||| |||  
417 GCACCAAGTCCGGAGCGGAGGATCCCGTCAAGACGCTCGACCGCGA 466  
51 u.....GlyProAlaAlaArgProAlaGlnGlnGlnL 63  
: |||  
467 CGCGCGGAGGCGCGCGCGCTCCCGCGCGCGAGGAGGAGGAGCGCG 516  
63 euAlaProLeuValProGlyPheLeuAlaProProPro..... 76  
|||||  
517 CCGGTCCGCGCGGTATGACGGCATG.....CCGGTGAACGGC 554  
77 .....Pr 77  
555 GAGAACAAATCTACCGCGCGCGCGCGGCGGCGGCTAAAGACAGCGGGTGC 604  
77 oAlaProAlaGlnSerProAlaPro...ThrGlnProProLeuProAspA 93  
|||||  
605 CACGCGCGCAGCGCGCGCGCTCCGTCGACCCAGAACAGAGCAGCGGTGA 654  
93 laglyValGlyGluLeu.....AlaProAspLeuLeuGlu 105  
||| |||  
655 ACGGTGAAACAAAGCTAAGCTCGCTCGCGCGCGCGCATAGCCGAG 704  
106 GlyIleAlaGluAspSerIleAspSerIleIleVal..... 117  
||| |||  
705 GCGCGCGCTCGGATTCGCGCGCTACCATTCATCAGCGACAAAGCGCC 754  
117 ..... 117  
755 GGAGTCCGTTGCCAGCTGAGAAAGACGCGCGCTCGTCCGGCTCAAAAT 804  
118 .....AlaAlaSerGluGlnAspSerGluIleMetAspAlaAsnGlu 131  
|||||  
805 TCGAGTCTCGCGCTCTGCTCCGCGGTCTGACACTGTCAGCGACGCGTGA 854  
132 Gln.....ProGlnAlaLy 136  
||| |||  
855 CAAGAAGTGAAGAGGCTGGGTGCTGTCGGAAGAAGCTCCAAAGCCAA 904  
136 sVal..... 137  
905 GGCTCTTCGCGCGCTGACGCGCGCTGTACAAAGAAGACCTTTGGGATT 954  
137 ..... 137  
955 TCAAGAAATACATTGTTTCGAGGAGCGCGCGGAGGAGGATGATGCG 1004  
137 ..... 137  
1005 CGGGCTGTCGAGATGATCGGGCTCTTTTGAACACCCAGGAGGATCAACA 1054  
138 .....ThrArgSerIleValPheValThrG 146  
||| |||  
1055 CTCCGAGACCTTTGGCAGGGGAGAGTGTATGACGCTGCTCGTGGCTG 1104  
146 LyGluAlaAlaProTyrAlaLysSerGlyLeuGlyAspValCysGly 162  
||| |||  
1105 CTGAGTGTCTCCCTGGTCAAAACAGGTGGTCTGGAGATGTTGGGGT 1154  
163 SerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMetValValMe 179  
|||||  
1155 GCTCTGCCCAAGGCTTTGGCAAGAGAGGACATCGTGTATGTTGTTGTT 1204  
179 tProArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLysAlaLeu 196  
|||||  
1205 ACCAAGGTAT.....GGGACTATGAAGAAGCCTACGATGTCG 1242



[illegible]

```

489 euGlySerGlyAspProIlePheGluGlyTrpMetArgSerThrGluSer 505
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
2104 TGGCGACCGCGCCACACACATCGAGAGCATGCTGCAGCATCTTCGACGG 2153
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
506 SerTrpLysAspLysPheArgGlyTrpValGlyPheSerValProValSe 522
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
2154 GAGCACACAGACAGGTGCGGGTGGGTGGGTCTCTCGTGCCTGGC 2203
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
522 rhiArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheG 539
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
2204 GCACGGATCATCGCGGGGGGAGACGCGCTCTCATGCGCTCCCGGTGC 2253
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
539 luProCysGlyLeuAsnGlnLeuTrpAlaMetGlnTrpGlyThrValPro 555
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
2254 AGCGTGGCGGCTGAACAGCTCTACGCCATGGCTACGGCACCGTCCCC 2303
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
556 ValValHisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnPr 572
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
2304 GTCGTGACCGCGTGGCGGCTCAGGGACACCGTGGCGCGCTTCGACCC 2353
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
572 oPheGlyAlaLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeut 589
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
2354 CTTC.....ACCACCTCCGGCTCGGTGGACGTTCGACCGCGCGCG 2394
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
589 hrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArg 605
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
2395 AGCGCACAGCTGATGACGAGCGCTCGGGCCTGCCTCCGCACCTACCGA 2444
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
606 GluHisLysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAs 622
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
2445 GACTTCAAGAGAGAGCTGGAGGGCCCTCCAGAGCGCGGCATGTGCGAGA 2494
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
622 phiThrTrpAspHisAla..ProSerSerThrSerArgSerSerSerGly 638
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
2495 CTTACGCTGGGAGCAGCGCGCCACAGCTCTACGAGGACGCTCTCGTCAAG 2544
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
639 ProSerTrpThrAsn.....ProThrSe 646
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
2545 CCAAGTACAGTGGTGAACGCTAGCTGCTAGCGCTCCAGCCCGCCATGC 2594
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
646 rCysargargGlyLeuGly.....ArgSerLysCysGlu...S 658
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
2595 GTGCATGACAGATGGAACTGCATTGGCAGCAGCAGGAAGTCCATGGAG 2644
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
658 erProSerAlaLeuLysThrSer 665
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
2645 CGCCGGCATCCCGAAGTACAGT 2667
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||:

seq_name: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:US-10-018-418-5

seq_document_block:
; Sequence 5, Application US/10018418
; GENERAL INFORMATION:
; APPLICANT: Morell, Matthew
; APPLICANT: Li, Zhongyi
; APPLICANT: Rahman, Sadequr
; APPLICANT: Appels, Rudolph
; TITLE OF INVENTION: Genes Encoding Wheat starch Syntheses and
; FILE REFERENCE: 127-01
; CURRENT APPLICATION NUMBER: US/10/018,418
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: PCT AU00/00385
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: AU PQ0052/99
; PRIOR FILING DATE: 1999-04-29
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2107
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS

```



LOCATION: (1)..(1791)  
US-10-018-418-5

## alignment\_scores:

Quality: 1154.50 Length: 661  
Ratio: 2.723 Gaps: 20  
Percent Similarity: 64.145 Percent Identity: 40.998

## alignment\_block:

US-09-674-824-2 x US-10-018-418-5

Align seg 1/1 to: US-10-018-418-5 from: 1 to: 2107

```

37 ArgArgLeuAlaArgGlyArgTyrValAlaGluLeuSerArgGluGlyPr 53
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
88 CAAGAACTGAAGAGGGTCCGGTCTTGTGNA.....GAAGCTCC 128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
53 oAlaAlaArgProAlaGlnGlnGlnGlnLeuAlaProProLeuValProG 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129 A.....AAGCCAAGGCTCTTTCGCCGCTGCAGCCCG 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
70 ly.....PheLeuAlaProProPro 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
164 CTGTACAGACACCTTTGGGATTCAGAAATACATTGGTTTCGAGGAG 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
77 ProAlaProAlaGlnSerProAlaProThrGlnProProLeuProAspAl 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
214 CCCGTGGAGGCCAAG.....GATGA 233
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
93 aGlyValGlyGluLeuAlaProAspLeuLeuGluGlyLeAlaGluA 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
234 TGGC.....CGGCTGTCGACATG 253
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
110 spSerIleAspSerIleValAlaAlaSerGluGlnAspSerGluIle 126
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
254 AT.....GCGGCTCTTTGAACACACACCAANT 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
127 MetAspAlaAsnGluGlnProGlnAlaLysValThrArgSerIleValPh 143
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
283 CACGACTCCGACCTTTGGCAGGGGAGAAATGTC...ATGACGTGGTGT 329
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
143 eValThrGlyGluAlaAlaProTyrAlaLysSerGlyLeuGlyAspV 160
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
330 CGTGGCTGCTAGTGTCTCTCTGTCGCAAAACAGGTGGTCTGGAGATG 379
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
160 aCysGlySerLeuProIleAlaLeuAlaAlaArgGlyHisArgValMet 176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
380 TTGGGGTGTCTGCCAAGGCTTTGGCAAGAGAGGACATCGTGTATG 429
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
177 ValValMetProArgTyrLeuAsnGlySerSerAspLysAsnTyrAlaLy 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
430 GTTGTGGTACCAGGTAT.....GGGACTATGAAGAACCTAC 467
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
193 sAlaLeuTyrThrAlaLysHisIleLysIleProCysPheGlySerH 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
468 GGATGTCGGAGTCCGAAATACTACAG.....GCTGCTGGACAGGATA 511
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
210 sGluValThrPhePheHisGluTyrArgAspAsnValAspTrpValPhe 226
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
512 TGGAAAGTGAATATTTCATCGTATATCATGATGAGTGTGATTTGTCTTC 561
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
227 ValAspHisProSerTyr...HisArgProGlySerLeuTyrGlyAspAs 242
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
562 ATTGACCTCTCTCTCTCCACACCGGAGGAGGAGACATTATGGG..... 606
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
242 nPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeuLeuCystyrA 259
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
607 ....GGCAGCAGACAGGAAATATTGAAGCCCATGATTTTCTTCTGCAAG 652
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
259 laAlaCysGluAlaProLeuLeuLeuLeuGlyGlyTyrIleTyrGly 275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
653 CCGCTGTGGTGTCCATGCACGCTTCCATGCGCGGTGTCCCTTATGGG 702
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

276 Gln...AsnCysMetPheValValAsnAspTrpHisAlaSerLeuValPr 291
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
703 GATGAAATCTGGTGTATTGCAAAATGATGGACACGGCACCTCTGCGC 752
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
291 oValLeuLeuAlaAlaLysTyrArgProTyrGlyValTyrArgAspSerA 308
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
753 TGTCTATCTGAAGCATATTACAGGGACCATGGTTGTATGCAGTACACTC 802
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
308 rGSerThrLeuValIleHisAsnLeuAlaHisGlnGlnValValGluProAla 324
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
803 GTCCATTATGTTGATACATAACATCGCTCACCAGGGCCGTGGCCCTGTA 852
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
325 SerThrTyrProAspLeuGlyLeuProProGluTyrTrpTyrGlyAlaLeuGl 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
853 GATGAATTCCTTCACCGAGTTGCTGAGCACTACCTG..... 891
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
341 uTrpValPheProGluTrpAlaArgArgHisAlaLeuAspLysGlyGluA 358
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
892 .....GAACACTTCAGACTGTACGACCCCGTGGTGGTGAAC 928
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
358 laValAsnPheLeuLysGlyAlaValValThrAlaAspArgIleValThr 374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
929 ACGCAAACTACTTCGCGCGGCTGAAGATGGCGGACCAAGTGTGCTGTG 978
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
375 ValSerGlnGlyTyrSerTrpGluValThrAlaGluGlyGlnGln 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
979 GTGAGCCCCGGGTACCTGTGGAGCTGAAGACGCTGGAGGGCGCTGGG 1028
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
391 yLeuAsnGluLeuLeuSerSerArgLysSerValLeuAsnGlyIleValA 408
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1029 GCTTCAGACATCATACGCGAGACAGCTGGAAGACCCGCGGCATCGTCA 1078
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
408 sNGLyIleAspIleAsnAspTrpAsnProThrThrAspLysCysLeuPro 424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1079 ACGSCATCGACAACATGGAGTGGAAACCCGAGGTGGACGCCACCTCAAG 1128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
425 His.....HisTyrSerValAspLeu...SerGlyLysAl 436
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1129 TCGACCGCTACACCACTCTCTCTGAGGACGCTGGACTCCGGCAAGCG 1178
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
436 aLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArgGluA 453
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1179 GCAGTGAAGGAGGCGCCCTGCAGCGAGCTGGCCCTGCAGGTCGCGCGCG 1228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
453 spValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIle 469
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1229 ACGTGGCGCTGCTCGGCTTCATCGCGCGCTGGACGGCAGAGGCGGTG 1278
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
470 AspleuIleLysMetAlaIleProGluLeuMetArgGluAspValGlnPh 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1279 GAGATCATCGCGACGCCATCGCTGATCGTGAGCCAGGACGTCGAGCT 1328
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
486 eValMetLeuGlySerGlyAspProPheGluGlyTyrPheMetArgSerT 503
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1329 GGTGTGCTGGGACCGCGCCACGACCTGGAGAGCATGCTGCAGCACT 1378
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
503 hrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerVal 519
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1379 TCAGCGGGGAGCACCGACAGCATGTCGCGGGTGGGTGGGTTCCTCCGTG 1428
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
520 ProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSe 536
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1429 CGCCTGGCGCACCGGATCAGCGGGGGGCGGACGCGCTCTCTCATGCCCTC 1478
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
536 rArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyT 553
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1479 CCGGCTGCTGGCGTGGCGGTGAACACAGCTCTACGCCATGGCCTACGGCA 1528
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
553 hrValProValValHisGlyThrGlyLeuArgAspThrValGluThr 569
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1529 CCGTCCCGCTGTCGACCGCGGTGGCGGCTCAGGGACACCGGTGGCGCGG 1578
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

570 PheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPheSe 586
|||||
1579 TTGACCCCTTC.....AACCACTCCGGGCTCGGGTGGACGTTTCCA 1619
586 rProLeuThrValAspLysMetLeuTrpAlaLeuArgThrAlaMetSert 603
|||
1620 CGCGCCCGAGCGCACAAAGCTCATCGAGCGCTCGGGCACTGCCTCCGCA 1669
603 hrPheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMet 619
1670 CCTACCGAGACTTCAAGAGAGCTGGAGGGCCCTCCAGAGCGCGGCATG 1719
620 ThrLysAspHisThrTrpAspHisAla_ProSerSerThrSerArgSers 636
|||
1720 TCGCAGGACTTCAGCTGGAGCAGCCGCCCAAGCTCTACGAGGAGCTCT 1769
636 exSerGlyProSerTrpThrAn..... 643
1770 CGTCAAGGCCCAAGTACCAGTGGTGAACGCTAGCTGCTAGCCGCTCCAGCC 1819
644 ProThrSerCysArgArgGlyLeuGly.....ArgSerLysCy 656
1820 CGCGATGGTGCATCAGAGTGGAACTGCATTCGCCGACGCGAGGAAGTG 1869
656 sglu...SerProSerAlaLeuLysThrSer 665
1870 CCATGGAGCGCGCATCGCGAAGTACAGT 1900

```

seq\_name: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:US-10-044-543-5

```

seq_documentation_block:
; Sequence 5, Application US/10044543
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: Novel Starch Synthase Polynucleotides
; TITLE OF INVENTION: and Their Use in the Production of New Starches
; FILE REFERENCE: 11440
; CURRENT APPLICATION NUMBER: US/10/044.543
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/388,743
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 2348
; TYPE: DNA
; ORGANISM: Curcuma zedoaria
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (36)...(2105)
US-10-044-543-5

```

```

alignment_scores:
  Quality: 1150.00      Length: 517
  Ratio: 3.125          Gaps: 13
  Percent Similarity: 71.180      Percent Identity: 47.389

```

alignment\_block:  
US-09-674-824-2 x US-10-044-543-5

Align seg 1/1 to: US-10-044-543-5 from: 1 to: 2348

```

140 SerIleValPheValThrGlyGluAlaAlaProTrpAlaLysSerGlyL 156
|||||
633 AATATCATATGCTAGCTGCAGAATCGCACCATGCTCTAAACACAGTGG 682
156 yLeuGlyAspValCysGlySerLeuProIleAlaLeuAlaAlaArgGlyH 173
|||||
683 GCTTGGAGATGTTGTTGGAGCTTTACCTAAAGCATTTGGCCAGAGAGGAC 732
173 isArgValMetValValMetProArgTyrLeuAsnGlySerSerAspLys 189

```

```

733 ATCGTGTTCATGGTAGTCTCCAGATAT.....GGA 764
190 AsnTyr.....AlaLysAlaLeuTyrThrAlaLysHisIleLysIlePr 204
|||||
765 AACTATCTGACCTAACCTAAGGAATAGGAATCTTAAAGGTTACAAGGTTGA 814
204 oCysPheGlyGlySerHisGluValThrPhePheHisGluTyrArgAspa 221
|||
815 T.....GGACAGCATCGAGATTAATACTATCATCTACATCATCGATT 856
221 snValAspTrpValPheValAspHisProSerTyrHisArgProGlySer 237
|||||
859 CTGTTGATTTGCTTCATCGATAGTCTATTTTCCGCCATATTTGGAAT 908
238 ..LeuTyrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTy 253
|||
909 GATATATATGTTGGAAC.....CGAGTGCACATTTTGAAGAGAAAT 949
253 rThrLeuLeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuG 270
950 GGTATTTCTTCAAGACAGCAGTTGAGTTCTTGGCATGTCCTCCCATGTG 999
270 yGlyTyrIleTyrGlyGln...AsnCysMetPheValValAsnAspTrp 285
|||||
1000 GTGATTTCTGTAAGAGATGGGAATTTGGTTTTCATTTGCCACGATTTGG 1049
286 HisAlaSerLeuValProValLeuAlaAlaLysTyrArgProTyrGl 302
|||||
1050 CATACCTCCTTACTTCCAGTTTATTGAAGGCATGTTTCGGTGATCGTGG 1099
302 yValTyrArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisG 319
1100 AATAATGACATACGCTCGCTCTCTTGGTTATTCAACAACATTCACATC 1149
319 lnGlyValGluProAlaSerThrTyrProAspLeuGlyLeuProProGlu 335
1150 AGGTTCGTGTCCTGCTAGTACTTCATATGTTGGATTTGCCACATGAT 1199
336 TrpTyrGlyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAl 352
1200 CACATTGACTCGTTTAGA..... 1217
352 aleuAspLys.....GlyGluAlaValAsnPheLeuLysGlyAlav 366
1218 .CTGGATGATCCTGTGGAGGTGAGCATTTTAACATTTTTCACAGCTGTA 1266
366 alValThrAlaAspArgIleValThrValSerGlnGlyTyrSerTrpGlu 382
1267 TAAGAGCTGCTGACCGTGTGTTACAGTTAGCCATGGCTATGCTTGGAG 1316
383 ValThrThrAlaGluGlyGlnGlyLeuAsnGluLeuLeuLeuSerSerAr 399
1317 TTAACAACTCTGAAGGTGGTGGGATTCATGAGATCATCAACGAGTG 1366
399 glysserValLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpA 416
1367 CCATTGGAAATTCATGTTATTTAAATGGAATCGATACCCATAGTTGGA 1416
416 snProThrThrAspLysCysLeuProHis.....HisTyrSer 428
1417 ATCCAAATTTGACGCTCACTTAAATTTCTGATGTTTACACCAACTTCACC 1466
429 ValAspAspLeuSer...GlyLysAlaLysCysLysAlaGluLeuGluLy 444
1467 CTGGAACACTTGAATGGGAAGGCCAGTGCAAGGCTGCTTTGCCAACG 1516
444 sGluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheIleG 461
1517 AGAGTTTGGTCTGCTGTTTCGTGACGAGCTTCTTATCTTGGCTTCATTG 1566
461 lyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIlePro 477

```



```

1567 GGAGATTAGACCACATCAAAAGGTATAGATCTCATAGCGAGGCCATGCAC 1616
478 GluLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAspPr 494
1617 TGGCTCGTCGCTCAGATCTACAGATAATATGCTGGGCACCTGGGAGGCC 1666
494 oilePheGluGlyTrpMetArgSerThrGluSerSerThrLysAspLysP 511
1667 AGACCTCGAGGATATGCTTCGAAGATTTCAGCTGAGCATCCGGTAAGG 1716
511 heArgGlyTrpValGlyPheSerValProValSerHisArgIleThrAla 527
1717 TCAGGGATGGGTGGGTCTCAGTGAATAATGGCTCATCGGATCAGCA 1766
528 GlyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLeuAs 544
1767 GGTGCTGATGCCCTACTATGCCCTCCAGGTTCGAACCTTTGGATTGAA 1816
544 nGlnLeuTyrAlaMetGlnTyrGlyThrValProValValHisGlyThrG 561
1817 CCAACTTCACGCTATGATGTACGGAAACAATTCCTGTGTGCATGCAGTAG 1866
561 lyGlyLeuArgAspThrValGlnThrPheAsnProPheGlyAlaLysGly 577
1867 GTGGTCTTCGAGATACTGTGCAACAGTTTGATCCGTTT.....AAT 1907
578 GluGluGlyThrGlyTrpAlaPheSerProLeuThrValAspLysMetLe 594
1908 GAGACAGCTTTGGATGGACCTTTGACAGGGCAGGCACATAGATGAT 1957
594 uTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisLysProSerT 611
1958 AGTGGCACTCGCCATTGCTCTAAACACATATCGGAATTACAAGGAGGCT 2007
611 rpGluGlyLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHis 627
2008 GGGTGGGATTTCAGAACGAGGGATGATGCAGGACCTCAGTTGGGAGAGT 2057
628 AlaProSerSerThrSerArgSerSerGlyProSerTrpThrAsn 643
2058 GCTGCCGAGCACTATGAAGAGTCTTGTCTGCTGCCAAGT...ACCAAT 2103
seq_name: /cgn2_6/ptodata/2/pna/us09_NEW_COMB_seq2:US-09-865-419A-41741
seq_documentation_block:
; Sequence 41741, Application US/09865419A
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51935)B
; CURRENT APPLICATION NUMBER: US/09/865,419A
; PRIOR FILING DATE: 2001-05-29
; PRIOR FILING DATE: 2000-05-31
; NUMBER OF SEQ ID NOS: 54020
; SEQ ID NO 41741
; LENGTH: 1936
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700100789_FLI
US-09-865-419A-41741

```

```

alignment_scores:
  Quality: 1068.50      Length: 467
  Ratio: 3.238          Gaps: 12
  Percent Similarity: 70.664  Percent Identity: 48.394
alignment_block:
US-09-674-824-2 x US-09-865-419A-41741

```

```

Align seg 1/1 to: US-09-865-419A-41741 from: 1 to: 1936
207 GlyGlySerHisGluValThrPhePheHisGluTyrArgAspAsnValas 223
59 GGACAGGATTTCAGAGTTTACTTATTTTTCACCTTACATTTGATGAGTTGA 108
223 pTrpValPheValAspHisProSerTyr...HisArgProGlySerLeuT 239
109 TTTTGTATTCTAGAACCCCTCCCTCCGCGCACCGGCACACATAATATTT 158
239 yrGlyAspAsnPheGlyAlaPheGlyAspAsnGlnPheArgTyrThrLeu 255
159 ATGGG.....GGAAAGATTGGATATATTTTGAAGCGCATGATTTTG 199
256 LeuCysTyrAlaAlaCysGluAlaProLeuIleLeuGluLeuGlyGlyTy 272
200 TTCTGCAAGCCGCTGTTGAGTTTCCATGGTATGCTCCATGTGGCGGTAC 249
272 rIleTyrGlyGln...AsnCysMetPheValValAsnAspTrpHisAlas 288
250 TGTCTATGGTGTGCAACTTAGTTTATTGCTAATGATTGGCATACCG 299
288 erLeuValProValLeuLeuAlaLysTyrArgProTyrGlyValTyr 304
300 CACTTCGCTGCTCTATCTAAAGCCCTATTACCGGGACAAATGGTTTGATG 349
305 ArgAspSerArgSerThrLeuValIleHisAsnLeuAlaHisGlnGlyVa 321
350 CAGTATGCTCGCTCTGCTGTGTGTATACACACATTCCTCATCAGGTCG 399
321 lGluProAlaSerThrTyrProAspLeuGlyLeuProGluTrpTyrG 338
400 TGCCCTGTAGAGCACTTCGTCAAATTTTGACTTGCCT..... 436
338 lyAlaLeuGluTrpValPheProGluTrpAlaArgArgHisAlaLeuAsp 354
437 .....GAACACTACATCCGAC 451
355 Lys.....GlyGluAlaValAsnPheLeuLy 363
452 CACTTCAAACTGATGACAACTTGTGGGGATCAGACCAACGTTTTCG 501
363 sGlyAlaValThrAlaAspArgIleValThrValSerGlnGlyTyrS 380
502 TCGGGGGCTGAAGACCGCAGACCGGGTGTGACCGTTAGCAATGGCTACA 551
380 erTrpGluValThrAlaGluGlyGlyGlnGlyLeuAsnGluLeuLeu 396
552 TGTGGGAGCTGAAGACTTCGGAAGCGGGTGGGGCTCCACGACATCATA 601
397 SerSerArgLysSerValLeuAsnGlyIleValAsnGlyIleAspIleAs 413
602 AACCAAGACGAGCTGGAAGCTGACGGGCATCGTGAACGGCATCGACATGAG 651
413 nAspTrpAsnProThrThrAspLysCysLeuProHis.....H 426
652 CGAGTGGAAACCCCGCTGTGGAGCTGCACCTCCACTCCGAGACTACACA 701
426 isTyrSerValAspAspLeu...SerGlyLysAlaLysCysLysAlaGlu 441
702 ACTACACGTTTCGAGACCGCTGGACACCGCAGCGCAGTGCAGGCGCC 751
442 LeuGlnLysGluLeuGlyLeuProValArgGluAspValProLeuIleG 458
752 CTGCAGCGGCGAGCTGGGCTTCAGGTCCGCGACGACGCTGCCACTGATCG 801
458 yPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetA 475
802 GTTCATCGGGGCTGGACCAACAGAGGGCGTGGACATCATCAGCCGACG 851
475 laileProGluLeuMetArgGluAspValGlnPheValMetLeuGlySer 491
|||||

```



US-09-674-824-2 x US-09-873-402A-83353  
Align seg 1/1 to: US-09-873-402A-83353 from: 1 to: 660  
369 AlaAspArgIleValThrValSerGlnGlyTyrSerTrpGluValThrTh 385  
3 GCAGATCGAATGCTGACTGCTCAGTAAGGGTTATTTCATGGGAGGTCAAC 52  
385 rAlaGluGlyGlnGlnGlyLeuAsnGlnLeuLeuSerSerArgLysSerV 402  
53 TGCTGAAGGTGGACAGGGGCTCAATGAGCTCTTAAGCTCCAGAAAGAGTG 102  
402 alLeuAsnGlyIleValAsnGlyIleAspIleAsnAspTrpAsnProThr 418  
103 TATTAAACGGAATTGTAATGGAATTGACATTAATGATTGGAACCTGCC 152  
419 ThrAspLysCysLeuProHisHisTyrSerValAspLeuSerGlyL 435  
153 ACAGACAAATGTATCCCTCTGCTTATTCTGTTGATGACCTCTCTGGA 202  
435 sAlaLysCysLysAlaGluLeuGlnLysGluLeuGlyLeuProValArg 452  
203 GGCAATATGTAAGGTGCATTGCAGAAAGGAGCTGGGTTTACCTATAAGGC 252  
452 luAspValProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGly 468  
253 CTGATGTTCTCTGATGGCTTTATTGGAAGATTGGATTATCAGAAGGC 302  
469 IleAspLeuIleLysMetAlaIleProGluLeuMetArgGluAspValG 485  
303 ATTGATCTCATTCACCTTATCATACACAGATCTCATCGGGGAAAGATTGTCA 352  
485 nPheValMetLeuGlySerGlyAspProIlePheGluGlyTrpMetArg 502  
353 ATTTGTCATGCTTGGATCTGGTGACCCAGAGCTTGAAGATTGGATGAGAT 402  
502 erThrGluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSer 518  
403 CTACAGATCGATCTTCAAGGATAAATTTCTGGATGGGTTGGATTAGT 452  
519 ValProValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetPr 535  
453 GTTCCAGTTTCCCACCGAATAACTGCCGCTGCCGATATATTGTTAAATGCC 502  
535 oSerArgPheGluProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrG 552  
503 ATCCAGATTCCGAACCTTGTGCTCTCAATCAGCTATATGCTATGCGAGTATG 552  
552 lyThrValProValValHisGlyThrGlyGlyLeuArgAspThrValGlu 568  
553 GCACAGTTCTCTGTTGTCATCCAACTGGGGCCCTTAGAGATACCGTGGAG 602  
569 ThrPheAsnProPheGlyAlaLysGlyGluGlyThrGlyTrpAlaPh 585  
603 AACTTCAACCCCTTTCCGTTGAGAATGGAGAGCGAGGTACAGGGTGGGCATT 652  
585 eSerPro 587  
653 CGCACCC 659

852 CGATCCACTGGATCGCGGGCAGAGACGTGACAGCTGCTGATGCTGGGCACC 901  
492 GlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSerTyrLy 508  
902 GGGCGGGCCGACCTGGAGGACATGCTCGCGGGTTCGAGTGGAGCACAG 951  
508 sAspLysPheArgGlyTrpValGlyPheSerValProValSerHisArgI 525  
952 CGACAAGGTGCGCGGTGGGTGGGTTCTCGGTGCCCTGGCGCACCGCA 1001  
525 leThrAlaGlyCysAspLeuLeuMetProSerArgPheGluProCys 541  
1002 TCACGGCGGGCGCGGACATCTGCTGATGCGCTGCGGTTCGAGCCGTGC 1051  
542 GlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValValH 558  
1052 GGGTGAACACAGCTCTAGCCCATGGCGTACGGGACCGTGGCGGTGCA 1101  
558 sGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheGlyA 575  
1102 CGCGGTGGGGGGTCCGGGACACGGTGGCGCTTCGACCCGTTCC... 1147  
575 laLysGlyGluGlyThrGlyTrpAlaPheSerProLeuThrValAsp 591  
1148 .....AACGACACCGGGTCCGGTGGAGCTTCGACCGCGCGGAGCGAAC 1192  
592 LysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisLy 608  
1193 CGGATGATCGAGCGCTCTCGCACTGCGCTACCACGTAACCGGAACACAA 1242  
608 sProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisThr 625  
1243 GGAGAGTGGCGCGCTCGACGGCGCGGATGGCGGAGGACCTCAGCT 1292  
625 rpAspHisAla.Pro.....SerSerThrSerArgSerSe 636  
1293 GGGACACCGCGCGCTGTGTATGAGGACGTCTGCTCAAGCGGAAGTAC 1342  
636 rSerGlyPro.....SerTrp...ThrAsnProThrSerCysArgArg 649  
1343 CAGTGGTGAGCGAATAATTTGGCGACGCGCGCGCTCTGCTGCCAGG 1391  
seq\_name: /cgn2\_6/ptodata/2/pns/US09\_NEW\_COMB.seq3:US-09-873-402A-83353  
seq\_documentation\_block:  
; Sequence 83353, Application US/09873402A  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; APPLICANT: Hardeman, Kristine J.  
; APPLICANT: Varagona, Marguerite J.  
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(51934)B  
; CURRENT APPLICATION NUMBER: US/09/873,402A  
; CURRENT FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/209,830  
; PRIOR FILING DATE: 2000-06-06  
; NUMBER OF SEQ ID NOS: 90966  
; SEQ ID NO 83353  
; LENGTH: 660  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3689-223-Q6-K6-E12  
US-09-873-402A-83353

alignment\_scores:  
Quality: 1062.00 Length: 219  
Ratio: 5.033 Gaps: 0  
Percent Similarity: 96.347 Percent Identity: 89.498  
alignment\_block:



---



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2002, 14:56:15 ; Search time 60.77 seconds  
(without alignments)  
947.638 Million cell updates/sec

Title: US-09-674-824-2  
Perfect score: 4044  
Sequence: 1 MAATGVGAGCLAPSRLRAD.....SDGSLSVRTAEIRNLQVTL 756

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3313	81.9	610	2 T06280	probable ADPglucos
2	2609.5	64.5	626	2 JQ2322	ADPglucose--starch
3	2591	64.1	622	2 T01414	ADPglucose--starch
4	1992	49.3	641	2 T07668	ADPglucose--starch
5	1227.5	30.4	732	2 T01208	ADPglucose--starch
6	1205.5	29.8	698	2 T01209	ADPglucose--starch
7	1204	29.8	752	2 S61505	ADPglucose--starch
8	1154.5	28.5	788	2 T07667	UDPglucose--starch
9	1128.5	27.9	491	2 T06798	UDPglucose--glycog
10	932.5	23.1	608	2 S43341	probable starch sy
11	884.5	21.9	607	1 YUPOY	ADPglucose--starch
12	883	21.8	608	2 T14731	ADPglucose--starch
13	879	21.7	609	2 JQ0703	UDPglucose--starch
14	877	21.7	609	2 S11481	UDPglucose--starch
15	876.5	21.7	603	2 S61504	UDPglucose--starch
16	870	21.5	605	1 S07314	UDPglucose--starch
17	863	21.3	610	2 F86453	granule-bound star
18	861	21.3	608	2 T10906	ADPglucose--starch
19	858.5	21.2	603	1 YUBHY	UDPglucose--starch
20	843.5	20.9	615	1 YUWTV	UDPglucose--starch
21	762.5	18.9	483	2 C86712	UDPglucose--starch
22	758	18.7	322	2 T07924	ADPglucose--starch
23	756	18.7	484	2 S40051	probable starch sy
24	703.5	17.4	486	2 H72321	ADPglucose--starch
25	685.5	17.0	477	1 SYECGL	glycogen synthase
26	685.5	17.0	477	2 C86009	ADPglucose--starch
27	654.5	16.2	476	2 E83785	glycogen synthase
28	645.5	16.0	484	2 F82165	starch (bacterial
29	626.5	15.5	476	2 C64119	glycogen synthase
					ADPglucose--starch

30	592.5	14.7	477	2 S76496	hypothetical prote
31	590.5	14.6	1071	2 T04926	starch synthase ho
32	589	14.6	463	2 C70363	glycogen synthase
33	579.5	14.3	472	2 B75501	glycogen synthase
34	563	13.9	491	2 S74473	probable ADPglucos
35	556	13.7	1230	2 T07663	soluble starch syn
36	535	13.2	1025	2 H86250	hypothetical prote
37	527.5	13.0	513	2 H83375	probable glycogen
38	509	12.6	441	2 T07926	probable starch sy
39	492	12.2	1674	2 T01265	starch synthase DU
40	419	10.4	474	2 D71469	probable glycogen
41	417	10.3	474	2 A81732	glycogen synthase
42	389.5	9.6	476	2 C72016	glycogen synthase
43	389.5	9.6	476	2 B86609	glycogen synthase
44	346	8.6	521	2 E64500	UDPglucose--starch
45	318.5	7.9	310	2 T01266	starch synthase DU

## ALIGNMENTS

RESULT 1

T06280  
probable ADPglucose--starch glucosyltransferase (EC 2.4.1.21) precursor - wheat (frag  
N:Alternate names: starch synthase  
C:Species: Triticum aestivum (common wheat)  
C>Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
C:Accession: T06280  
R:Block, M.; Loerz, H.; Luettticke, S.  
submitted to the EMBL Data Library, February 1996  
A:Reference number: 215595  
A:Accession: T06280  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-610 <BLO>  
A:Cross-references: EMBL:U48227; NID:gl373149; PIDN:AAB02197.1; PID:gl373150  
A:Experimental source: cv. Florida, endosperm of kernels  
C:Function:  
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producin  
C:Keywords: glycogen/starch biosynthesis; glucosyltransferase; hexosyltransferase

Query Match 81.9%; Score 3313; DB 2; Length 610;  
Best Local Similarity 100.0%; Pred. No. 8.1e-328;  
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 147 EAAPYAKSGGLDVCGLPIALAAARGHRVVMVMPRYLNGSSDKNYAKALYTAKHIPCF 206

Db 1 EAAPYAKSGGLDVCGLPIALAAARGHRVVMVMPRYLNGSSDKNYAKALYTAKHIPCF 60

Qy 207 GGSHEVTFFEHYRDNDVAVFDVDPHSYHRPGSLYGDNFGAFGDNQFRYTLACYACEAPLI 266

Db 61 GGSHEVTFFEHYRDNDVAVFDVDPHSYHRPGSLYGDNFGAFGDNQFRYTLACYACEAPLI 120

Qy 267 LELGGYIYGQCMFVYNDVHSLVPVLLAAKRYPGYVYRDSRSLVTHNLAHQGVPEAST 326

Db 121 LELGGYIYGQCMFVYNDVHSLVPVLLAAKRYPGYVYRDSRSLVTHNLAHQGVPEAST 180

Qy 327 YPDGLPPEWYGALEWFPPEWARHLDKGEAVNFKGAVVTADRIVTVSOGYSWEVTTA 386

Db 181 YPDGLPPEWYGALEWFPPEWARHLDKGEAVNFKGAVVTADRIVTVSOGYSWEVTTA 240

Qy 387 EGGGGLNELLSSRSKSVLNGIVNGIDINDNPTTDKCLPHHYSDVDDLGGKAKCAELQKEL 446

Db 241 EGGGGLNELLSSRSKSVLNGIVNGIDINDNPTTDKCLPHHYSDVDDLGGKAKCAELQKEL 300

Qy 447 GLPVREDVPLIGTGRLDYQKIDIKMAIPELMREDVQFVMLGSGDPIFEQWMRSTESS 506

Db 301 GLPVREDVPLIGTGRLDYQKIDIKMAIPELMREDVQFVMLGSGDPIFEQWMRSTESS 360

Qy 507 YKDFRCWGVFSVPVSHRITAGCDILLMPSPFPCGLNQLYAMOYGTVPVYVHTGGGLRDT 566

Db 361 YKDFRCWGVFSVPVSHRITAGCDILLMPSPFPCGLNQLYAMOYGTVPVYVHTGGGLRDT 420



567 VETNFPGAKGEGCTGAWPSPLTVDKMLALRTAMSTFFREHKPSWEGLMKRGMTKDHTWD 626  
421 VETNFPGAKGEGCTGAWPSPLTVDKMLALRTAMSTFFREHKPSWEGLMKRGMTKDHTWD 480  
627 HAPSSSTRSSSGSWTNPSTSCRRGLGRSKCESPSALKTSSTSSSFRGPEGYPCTLRCPATVE 686  
481 HAPSSSTRSSSGSWTNPSTSCRRGLGRSKCESPSALKTSSTSSSFRGPEGYPCTLRCPATVE 540  
687 SOCACLLWFAGSRTYDGCRAAAAVTASGGRLQFQWGRKCAAGWLTAKHHSOGSLSVRYT 746  
541 SOCACLLWFAGSRTYDGCRAAAAVTASGGRLQFQWGRKCAAGWLTAKHHSOGSLSVRYT 600  
747 AEIRNQLVTL 756  
601 AEIRNQLVTL 610  
RESULT 2  
JQ2322  
ADPglucose--starch glucosyltransferase (EC 2.4.1.21) precursor - rice  
N:Alternate names: starch synthase  
C:Species: Oryza sativa (rice)  
C:Accession: JQ2322; PQ0811  
R:Baba, T.; Nishihara, M.; Mizuno, K.; Kawasaki, T.; Shimada, H.; Kobayashi, E.; Ohnishi  
Plant Physiol. 103, 565-573, 1993  
A:Title: Identification, cDNA cloning, and gene expression of soluble starch synthase in  
A:Reference number: JQ2322; MUID:94302151  
A:Accession: JQ2322  
A:Molecule type: mRNA  
A:Residues: 1-626 <BAB1>  
A:Cross-references: DDBJ:D16202; NID:g450484; PIDN:BAA03739.1; PID:g450485  
A:Accession: PQ0811  
A:Molecule type: protein  
A:Residues: 114-129 <BA2>  
A:Experimental source: seed  
C:Function:  
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producing e  
C:Superfamily: starch synthase  
C:Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase  
F:1-113/Domain: transit peptide (amyloplast) #status predicted <TRS>  
F:114-626/Product: ADPglucose--starch glucosyltransferase, 57kd form #status experimental  
F:122-626/Product: ADPglucose--starch glucosyltransferase, 55kd form #status experimental  
F:440-448/Region: substrate binding #status predicted

Query Match 64.5%; Score 2609.5; DB 2; Length 626;  
Best Local Similarity 82.6%; Pred. No. 9.5e-178;  
Matches 497; Conservative 32; Mismatches 42; Indels 31; Gaps 9;  
2 AATGVGAGCL-APSVRLRADPATAARASACVVRARLRRLRGRVVAELSREGPAARPAQ 60  
5 AGMGIGAACLVAQVR-----PGR-----RLRLQVR--RCVAELSREGGSA----H 46  
61 QOLAP-----PLVPGFLAP--PPAPAQS--PAPTQPLPDAGVGLAPDLLLEGTAED 110  
47 GPLAPVLKQVLPFLVPTSTTAPQSPAPAPPTPPLPDSGVGEIEPD--LEGLETD 104  
111 STDSIIVAASEODSEIMDANEQPAKVTTSIVFTVGEAARYAKSGGLDVCGLSPLALAA 170  
105 SIDKTIIVASEGESEIMDVKEQAQAKVTSVVFVTEASPYAKSGGLDVCGLSPLALAL 164  
171 RGRHVMVMPRYLNGSSDKNYAKALYTAHKIKIPCFGSGHEVTFPHEYRDNDVWVVDHP 230  
165 RGRHVMVMPRYMGNALKNFANAFYTEKHKIPCFGSGHEVTFPHEYRDNDVWVVDHP 224  
231 SYHRGSLYDNGFAGDQNFRTLLCYAAACEAPLILELGGYIYGQNCMEFVNDWHASLV 290  
225 SYHRGSLYDNGFAGDQNFRTLLCYAAACEAPLILELGGYIYGQNCMEFVNDWHASLV 284  
291 PYLLAAKRYPRYGVYRDSRSTLIHNLAHQGVPEPASTYPDGLPPENYGALEWVPEWARR 350

285 PYLLAAKRYPRYGVYRDSRSTLIHNLAHQGVPEPASTYPDGLPPENYGALEWVPEWARR 344  
351 HALDKGEAVNFKGAVVTADRTVTSQGYSWEVTTAEGGQGLNELLSSRKSVLNGVNGI 410  
345 HALDKGEAVNFKGAVVTADRTVTSQGYSWEVTTAEGGQGLNELLSSRKSVLNGVNGI 404  
411 DINDWNPSTDKCLPHHYSYDDLSGKAKCAELQKELGLPVREDYPLIGFICRLDYQKID 470  
405 DINDWNPSTDKCLPHHYSYDDLSGKAKCAELQKELGLPVREDYPLIGFICRLDYQKID 464  
471 LTKMAIPELMRDVOFVMLGSDPIFEQWMRSTESYKDKFGRVWGVFPVPSYSHRTAGCD 530  
465 LTKMAIPDLMRDNIQFVMLGSDPIFEQWMRSTESYKDKFGRVWGVFPVPSYSHRTAGCD 524  
531 ILLMPSREPCGLNQLYAMQYGTVPVHGTGGLRDTVETFPFGAKGEGTGWAFSPLTV 590  
525 ILLMPSREPCGLNQLYAMQYGTVPVHGTGGLRDTVETFPFGAKGEGTGWAFSPLTV 584  
591 DK 592  
585 EK 586  
RESULT 3  
T01414  
ADPglucose--starch glucosyltransferase (EC 2.4.1.21) precursor - maize  
N:Alternate names: starch synthase I  
C:Species: Zea mays (maize)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 11-Jan-2000  
C:Accession: T01414  
R:Knight, M.E.; Harn, C.; Lilley, C.E.R.; Guan, H.P.; Singletary, G.W.; Mu-Forester,  
Plant J. 14, 613-622, 1998  
A:Title: Molecular cloning of starch synthase I from maize (W64A) endosperm and expre  
A:Reference number: Z14279; MUID:98340555  
A:Accession: T01414  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-622 <KN1>  
A:Cross-references: EMBL:AF036891; NID:g2828011; PIDN:AAB99957.1; PID:g2828012  
A:Experimental source: strain W64A; endosperm  
C:Genetics:  
A:Gene: Ssl  
C:Function:  
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producin  
C:Superfamily: starch synthase  
C:Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 64.1%; Score 2591; DB 2; Length 622;  
Best Local Similarity 83.0%; Pred. No. 1.9e-176;  
Matches 492; Conservative 29; Mismatches 58; Indels 14; Gaps 4;  
4 TGVGAGCLAPSVRLRADPATAARASACVVRARLRRLRGRVVAELSREGPAARPAQ 60  
5 SAVGACALL-----LARAHPAAVGDRAARRRRLRRCVLAELSREGPAARPLPP 56  
61 QOLAPVLPGFLAPPPPPAPAQSAPTQPLPDAGVGLAPDLLLEGTAEDSIIVAS 120  
57 ALLAPVLPGFLA--PPAEPTEGEPASTPPVPPDAGLDGLG--LEPEGIAEGSIDNTVVAS 113  
121 EQDSEIMDANEQPAKVTTSIVFTVGEAARYAKSGGLDVCGLSPLALAAARGHVMVMP 180  
114 EQDSEIVVGEQARAKVTQTSIVFTVGEAARYAKSGGLDVCGLSPLALAAARGHVMVMP 173  
181 RYLNSSDKNYAKALYTAHKIKIPCFGSGHEVTFPHEYRDNDVWVVDHPSPHYPGSLY 240  
174 RYLNSTSKNYANAFYTEKHKIPCFGSGHEVTFPHEYRDNDVWVVDHPSPHYPGSLY 233  
241 DNFGAFDQNFRTLLCYAAACEAPLILELGGYIYGQNCMEFVNDWHASLVLAAYKRP 300  
234 DKFGAFDQNFRTLLCYAAACEAPLILELGGYIYGQNCMEFVNDWHASLVLAAYKRP 293  
301 YGVYRDSRSTLIHNLAHQGVPEPASTYPDGLPPENYGALEWVPEWARRHLDKGEAVN 360



Db 294 YGVYKDSRSILVHNLAHQGVPEASTYDGLGLPEWGALEWFPPEWARRHALDKGAVN 353  
QY 361 FLKGAIVTADRIYVSGYSWEVYTAGGGGGLNELLSSRSKSVLNGIYNGIDINDWNPSTD 420  
Db 354 FLKGAIVTADRIYVSGYSWEVYTAGGGGGLNELLSSRSKSVLNGIYNGIDINDWNPSTD 413  
QY 421 KCLPHYSVDDLSKAKCKAELQELGLPVREDVPLIGFGRDLYOKGIDLIKMAPELM 480  
Db 414 KCIFCHYSVDDLSKAKCKAELQELGLPIRDPVPLIGFGRDLYOKGIDLIQIPLDLM 473  
QY 481 REDVQFVNLGSGDPIFGWMRSTESSYKDKFRGMVGFSPVSHRITAGCDILLMPSRFEP 540  
Db 474 REDVQFVNLGSGDPELEDNMRSTESIFKDKFRGMVGFSPVSHRITAGCDILLMPSRFEP 533  
QY 541 CGLNOLYAMOYGTVPVHGTGGGLRDVTETNPFCAKEEGTGNAFSPSLTVDK 593  
Db 534 CGLNOLYAMOYGTVPVHGTGGGLRDVTETNPFCAKEEGTGNAFSPSLTVDK 586  
RESULT 4  
T07668  
ADPglucose--starch glucosyltransferase (EC 2.4.1.21) SSI precursor - potato  
C:Species: Solanum tuberosum (potato)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Jun-2000  
C:Accession: T07668  
R:Abel, G.J.W.  
submitted to the EMBL Data Library, January 1997  
A:Reference number: Z16081  
A:Accession: T07668  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-641 <ABE>  
A:Cross-references: EMBL:X10416  
A:Experimental source: cv. Desiree; leaf  
C:Genetics:  
A:Gene: SSI  
C:Function:  
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producing  
A:Pathway: starch synthase  
C:Superfamily: starch synthase  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 49.3%; Score 1992; DB 2; Length 641;  
Best Local Similarity 67.5%; Pred. No. 8.4e-134;  
Matches 360; Conservative 80; Mismatches 87; Indels 6; Gaps 4;

QY 102 LLLEGIAEDS--IDSTIVAASQD--SETMDANEQPAQVTRSIYVVTGEAAPYAKSGGL 157  
Db 90 LIPHVAGDATWVESHDIVANDRDLSDETEMEETPIKLTNIIIFVTAEAAPYAKSGGL 149  
QY 158 GDVCGSLPITALAARGHRVVMVPRYLNGS-SDNKYAKALYTAKHKIPCFGSGHEVTFPH 216  
Db 150 GDVCGSLPITALAARGHRVVMVPRYLNGSPSDEKYANAVLDVDRATVHCFGDAQEVAHYH 209  
QY 217 EYRDNDVWVFDHPSPYHPSGLYGNFGAFGNQFRYLLCYAAEAPLILELGYIYQG 276  
Db 210 EYRAGVDWVFDHSSYCRPGTPIYDYGAFGNQFRYLLSHAAEAPLVLPLGFTYGE 269  
QY 277 NCMFVYNDWHSVLPVLLAAKRYPGVYRDRSRTSLVHNLAHQGVPEASTYDGLGLPEW 336  
Db 270 KCLFLANDWHAALVPLLLAAKRYPGVYKARSIVAIHNIAHQGVPEASTYDGLGLPEW 329  
QY 337 YGALEWFPPEWARRHALDKGEAVNFKGAVVTADRIYVSGYSWEVYTAGGGGGLNELL 396  
Db 330 YGAVEMIFPTWARHALDTGETVNVYKGAIAVADRILTVSGYSWEITTPGGYGLHELL 389  
QY 397 SSRKSVLNGIYNGIDINDWNPSTDKCLPHYSVDDLSKAKCKAELQELGLPVREDVPL 456  
Db 390 SSRQSVLNGIYNGIDINDWNPSTDHIAHSYINDLSGKVGQCKTDLQELGLPIRDPDCL 449  
QY 457 IGFGRDLYOKGIDLIKMAPELMREDVQFVNLGSGDPIFGWMRSTESSYKDKFRGMV 516

Db 450 IGFGRDLYOKGIDLIKMAPELMQNDVQVVMILGSGEKQYEDWNRHTENLFDKFRWVG 509  
QY 517 FSVPSVSHRITAGCDILLMPSRFEPCCGLNOLYAMOYGTVPVHGTGGGLRDVTETNPFCAK 576  
Db 510 FSVPSVSHRITAGCDILLMPSRFEPCCGLNOLYAMRYGTIPVHSTGGGLRDVTDFNYPYAOE 569  
QY 577 G-EETGWAFFSPLTVDKMLMALWALTAMSTPREHKPSWEGLMKRGMTKDHTWDHA 628  
Db 570 GIGEGTGWTFSPLTSEKLLDCLKLAIGTYTEHKSSWEGLMRGMGRDYSWENA 622

## RESULT 5

T01208  
ADPglucose--starch glucosyltransferase (EC 2.4.1.21) isoform STSII-1 - maize (fragmen  
N:Alternate names: starch synthase isoform STSII-1  
C:Species: Zea mays (maize)  
C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 17-Nov-2000  
C:Accession: T01208  
R:Knight, M.E.; Hahn, C.; Lilley, C.E.R.; Guan, H.P.; Singletary, G.W.; Mu-Forester,  
Plant J. 14, 613-622, 1998  
A:Title: Molecular cloning of starch synthase I from maize (W64A) endosperm and expre  
A:Reference number: Z14279; MUID:98340555  
A:Accession: T01208  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-732 <KNI>  
A:Cross-references: EMBL:AF019296; NID:g2811133; PIDN:AAD13341.1; PID:g2655029  
A:Experimental source: strain W64A; endosperm  
C:Genetics:  
A:Gene: SSI1a  
C:Function:  
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose produc  
A:Superfamily: starch synthase  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 30.4%; Score 1227.5; DB 2; Length 732;  
Best Local Similarity 40.3%; Pred. No. 2.5e-79;  
Matches 276; Conservative 104; Mismatches 212; Indels 93; Gaps 16;

QY 12 APSVRLRAD-----PATAAARASACVVRRLRLRLAR-----GRY---VAELSREGPA 54  
Db 58 AALYRAEAEGKDPAPERSGDAARLPARRNAVSKRDPLOQVGRYGSATGNTARTGAA 117  
QY 55 ARPAQOQOLAPLVPGLFAPPPAPAPQSPAP-TOPPLPDAGVGLAPDLLL----- 104  
Db 118 S--CONAALADVEIKSIVAAPTSTIVKFPAPGYRMILPS---GDIAPETVLPAPKPLHES 172  
QY 105 -----EGIAEDSIDSTIVAASEQDSSEIMDANEQPAQK----- 136  
Db 173 PAVDGDSDNGIAPPIVEPLVQEAATWDFKYYIGFDEPDEAKDDSRVGCADDAGSFEHYGND 232  
QY 137 -----VTRTSIVFTVGEAAYAKSGGLDVCGLSLPIALAAARGHRVVMVPRYLNGSSDKN 190  
Db 233 GPLAGENVVMVIVAAECSPWCKTGGDGVVYGALPKALARGRHVRVMPVRY----- 286  
QY 191 YAKALYTAKHKIPCFGSGHEVTFEYRDNDVWVFDHPSPY-HRPGSLYGDNFGAFGDN 249  
Db 287 YVEAFDMGIRKYKKAAGQDLEVNRYFHAFIDGVDFVFIADPLFRHQDDIYG---GSRQEI 343  
QY 250 QFRTLLCYAAEAPLILELGGYIYQG-NCMFVYNDWHSVLPVLLAAKRYPGVYRDRS 308  
Db 344 MKRMILCKVAVEPWHVPCGVCYGDGNLYFIANDHHTALLPYLYKAYRDHGLMQVTR 403  
QY 309 STVIHNLHAQGVPEASTYDGLGLPEWYGALEWFPPEWARRHALDKGEAVNFKGAVVT 368  
Db 404 SVLVHINIAHQGRGPVDEFPYMDLPEHYLQHFELYDPV-----GGEHANIFAAGLKM 455  
QY 369 ADRTVTVSGYSWEVYTAGGGGGLNELLSSRSKSVLNGIYNGIDINDWNPSTDKCLPH--- 425  
Db 456 ADRTVTVSGYSWEVYTAGGGGGLNELLSSRSKSVLNGIYNGIDINDWNPSTDKCLPH--- 425







QY 399 RKSLVNGIVNGIDINDWNPDTDKCLPH-----HYSVDDL-SGKAKCAELQKELGLPVRED 453  
 Db 506 SDWKFRGIVNGYDTDKWNPQDAYTSDGYTNINLTKTQTKRQCKAALQRELGLPVRED 565  
 QY 454 VPLIGFGRDLVQKIDILKMAIPELMREDYQFVLMGSDPIFFGWMRSTESSYKDKFRG 513  
 Db 566 VPIISFGRDLHOKGVDLIAEAIPMMSHDVQVLMGTGRADLQMLKEFAQCHDKIRS 625  
 QY 514 WYGFVSPVSHRTTAGCDTLLMPSRFEPCGLNOLYAMQYGVVPHVHGCGGLRDTVETENPF 573  
 Db 626 WYGFVSKAHRITAGSDTLLMPSRFEPCGLNOLYAMSYGVVPHVHGCGGLRDTVQENPF 685  
 QY 574 GAKGEGTGWAFSPITVDKMLWALRTAMSTPREHKPSWEGLMKRGMTKDHWDHA 628  
 Db 686 ---DESGVGWTFDRAEANKLMAALWNCLLTYKYDKKSWEGIQERMSQDLSWDNA 737  
 RESULT 8  
 T06767  
 UDPglucose--glycogen glucosyltransferase (EC 2.4.1.11) precursor (clone GT11) - potato  
 N:Alternate names: glycogen (starch) synthase  
 C:Species: Solanum tuberosum (potato)  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
 C:Accession: T07667  
 R:Edwards, E.A.; Marshall, J.; Sidebottom, C.; Visser, R.G.F.; Smith, A.M.; Martin, C.  
 Plant J. 8, 283-294, 1995  
 A:Title: Biochemical and molecular characterisation of a novel starch synthase from potato  
 A:Reference number: Z16079; MUID:95400340  
 A:Accession: T07667  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-788 <EDW>  
 A:Cross-references: EMBL:X87988; NID:g887648; PIDN:CAA61241.1; PID:g887649  
 A:Experimental source: cv. Desiree; tuber  
 C:Comment: This protein is present both in the stroma of the plastid and tightly bound to  
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 28.5%; Score 1154.5; DB 2; Length 788;  
 Best Local Similarity 43.7%; Pred. No. 4.3e-74;  
 Matches 253; Conservative 83; Mismatches 194; Indels 49; Gaps 12;

QY 79 PAOSPAFTQPL-----PDAGVCELAPDILLLEGIA-----EDSDSIIVAAS 120  
 Db 215 PPETPKSSQETLLDVNSRSLVDVPGKKIQSYMPSLRKSSSHVQORNELEGSSAEAN 274  
 QY 121 EQDSEINDANEQPAQV---TRSVFTGTEAAPYAKSGGLGDCVGLPDLAALARGHRVW 177  
 Db 275 EETEDPVNIDEKPPPLAGTNVMMIILVASECAPWSKTGGLDVAGALPKALARRGRVW 334  
 QY 178 VMPRYLNGSSDKNKAALYAKHIKIPCGSGSHEVTFEYRDNDVWVTV-DHPSYHRPG 236  
 Db 335 VAPRYDNYPEQDSG---VRKIYKVD--GQDQVDTYFOALLMDCDFVFIHSHMFRHIGN 388  
 QY 237 SLYGDNFGAFDQNFRTLLCYAACAPLILLEGYIYGQ-NCMFVNDWASHVPLVLA 295  
 Db 389 NIYGGN---RVDILKRWLFCKAAIEVPWHPVCGVCYGDGNLVFTANDWHTALLPAYLK 445  
 QY 296 AKYRPYGVYRDSRSTLVHNLHQGVPEASTYDPLGLPPEWYGALEWVPEWARRHALDK 355  
 Db 446 AYRDNGICINNYTRSLVLIHNIHQGRGPLEDSYVDLPHPYMDPFKLYDPV-----G 497  
 QY 356 GEAVNFKGAVVTADRIIVTVSOGYSWEVTTARGGGLNELLSSRSKSVLNGIDINDW 415  
 Db 498 GEHNFIFAAGLKTADRVTVTVSHGYSWELKTSQGGNGLHQIINENDWKLOGIVNGIDTKEW 557  
 QY 416 NPTTDKCLPH-----HYSVDDL-SGKAKCAELQKELGLPVREDYPLIGFGRDLVQKID 469  
 Db 558 NPEDVHLPRSDGYNYSLSLDTLQTKPCKAALQELGLPVREDYPLIGFGRDLVQKID 617  
 QY 470 DLKMAIPELMREDYQFVLMGSDPIFFGWMRSTESSYKDKFRGWGFSVPVSHRITAGC 529  
 Db 618 DLIAEAVPMWQDQVQLVNLGTGRDRDLQMLQKQFECQNDKIRGWGFSVKTSHRITAGA 677

QY 530 DILLMPSRFEPCGLNOLYAMQYGVVPHVHCTGGRLDRTVETNPFGAKGEGTGWAFSPLT 589  
 Db 678 DILLMPSRFEPCALNOLYAMKYGTIPVHAVGRLDRTVQFPDPLMS---QDMGSPDRAE 734  
 QY 590 VDKMLWALRTAMSTPREHKPSWEGLMKRGMTKDHWDHA 628  
 Db 735 ASQLPIRINCLLTYREYKKSWEGIQTRCMTQDLSWDNA 773  
 RESULT 9  
 T06798  
 probable starch synthase (EC 2.4.1.1) - wheat (fragment)  
 C:Species: Triticum aestivum (common wheat)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 03-Nov-2000  
 C:Accession: T06798  
 R:Walter, L.; Loerz, H.; Lueticke, S.T.  
 submitted to the EMBL Data Library, August 1996  
 A:Reference number: Z15823  
 A:Accession: T06798  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-491 <WAL>  
 A:Cross-references: EMBL:U66377; NID:gl620659; PIDN:AAB17085.1; PID:gl620660  
 A:Experimental source: cv. Florida; endosperm; 21 days after anthesis  
 C:Genetics: SSI  
 A:Gene: SSI  
 C:Superfamily: starch synthase  
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 27.9%; Score 1128.5; DB 2; Length 491;  
 Best Local Similarity 47.6%; Pred. No. 1.6e-72;  
 Matches 236; Conservative 70; Mismatches 163; Indels 27; Gaps 9;

QY 140 SIVFVTGEAAPYAKSGGLGDCVGLPDLAALARGHRVWVMPVRYLNGSSDKNKAALYAK 199  
 Db 1 NVVVAAECSPCKTGGGLDVAGALPKALAKRGHRVWVMPVRYLNGSSDKNKAALYAK 56  
 QY 200 HKIPCFGSGSHEVTFEYRDNDVWVTV-DHPSY-HRPGSLYGDNFGAFDQNFRTLLCY 258  
 Db 57 YTK---AAGQDMENVYFHAYIDGVDFVTDAPLFRHQEDYIG---GSRQELMKRMILFK 111  
 QY 259 AACEAPLILELGGYIYGQ-NCMFVNDWASHVPLVLAALARGHRVWVMPVRYLNGSSDKNKAALYAK 317  
 Db 112 AAVEVPHVPCGGVPGDGNLVFTANDWHTALLPYLKAYYRDHGLMQYTRSIMVHNIA 171  
 QY 318 HGVPEASTYDPLGLPPEWYGALEWVPEWARRHALDKGEAVNFKGAVVTADRIIVTVSQ 377  
 Db 172 HOGRPVDFEFPTELPPEHYL-----EHPRLYDPVSGEHAHYFAAGLKADQVWVVP 223  
 QY 378 GYSWEVTTAEGGGLNELLSSRSKSVLNGIDINDWNPDTDKCLPH-----HYSVDDL- 432  
 Db 224 GYLWELKTVGEGWGLHDIIRQNDWKTRGIVNGIDNMENNEPEVDALHKSQDGTNFSRLTLD 283  
 QY 433 SGKAKCAELQKELGLPVREDYPLIGFGRDLVQKIDILKMAIPELMREDYQFVLMGSG 492  
 Db 284 SGKQCKEALQRELGLQVRAADVPLLGLFGRDLGQKGVETIADAMPWISQDVQLVMLGTG 343  
 QY 493 DPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDILMPSRFEPCGLNOLYAMQY 552  
 Db 344 RHDLESMLQHFERGHDKVRGWGFSVRLAHRITAGADAKLMPSRFEPCGLNOLYAMAYG 403  
 QY 553 TVPVVHGTGGLRDTVETNPFGAKGEGTGWAFSPLTVDKMLWALRTAMSTFREHKPSWE 612  
 Db 404 TVPVVHAGGLRDTVPPDFE---NHSGLGWTFDRAEHLKLEALHCLRTYRDFKESWR 460  
 QY 613 GLMKRGMTKDHWDHA 628  
 Db 461 ALQERMSQDSWEHA 476  
 RESULT 10



```

RESULT 11
YUPOY
ADPglucose--starch glucosyltransferase (EC 2.4.1.21) precursor - potato
N:Alternate names: starch synthase
C:Species: Solanum tuberosum (potato)
C:date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 11-Jun-
C:Accession: S16555; S24392; S26060; S26061

```







Db 311 GILEADRVLTSPYYAEELISCIARGCELDNIM--RLTGITIVNGMDVSEWDSKDKYI 368  
Qy 424 PHHY-SVDDLSCKAKCAELOKELGCLPVREDVPLIGFGRGLDYQKGLDKLIMKALPELMRE 482  
Db 369 TAKYDATTATIEAKALKALQAGLPVDRKIPLIATIGRLEEQKQPDVMAAAIPELMQE 428  
Qy 483 DVQFVMLGSGDPIFGWMRSTESSYKDKFRGWGVFSPVSHRITAGCDILLMPSRFPCG 542  
Db 429 DVQIVLLGTGKKFKELKSMEEKYPGKVRVAVKFNAPLAHLIMAGADVLAVPSRFPCG 488  
Qy 543 LNQLYAMQYGVTVVHGTCGLRDIVETNPFCAKGEECTGNAFSLTVVD-----K 592  
Db 489 LIQLQGMRYGTPCACASTGGLVDTVI-----EGTGFHMGRLSDVCKVPEPSDYKK 539  
Qy 593 MLWALRTAMSTPREHKPSWEGMLKRGMTKDHTWDHAPSSSTRSSSGP--SWTN 643  
Db 540 VAATLKRAIKVV--GTPAYEEMVRNQMQLSW-----KGPKNWEN 579  
RESULT 15  
S61504  
UDPglucose--starch glucosyltransferase (EC 2.4.1.11) Isoform I precursor - garden pea  
N:Alternate names: glycogen(starch) synthase  
C:Species: Pisum sativum (garden pea)  
C:Date: 20-Jul-1996 #sequence\_revision 07-Feb-1997 #text\_change 18-Jun-1999  
C:Accession: S61504; S72372  
R:Dry, I.; Smith, A.; Edwards, A.; Bhattacharyya, M.; Dunn, P.; Martin, C.  
Plant J. 2, 193-202, 1992  
A:Title: Characterization of cDNAs encoding two isoforms of granule-bound starch synt  
A:Reference number: S61504; MUID:93251108  
A:Accession: S61504  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-603 <DRY>  
A:Cross-references: EMBL:X88789; NID:g887570; PIDN:CAA61268.1; PID:g887571  
A:Accession: S72372  
A:Molecule type: protein  
A:Residues: 76-77, X', 79, 81-88 <DRW>  
C:Superfamily: starch synthase  
C:Keywords: glucosyltransferase; hexosyltransferase  
F:1-75/Domain: signal sequence #status predicted <SIG>  
F:76-603/Product: glycogen (starch) synthase Isoform I #status experimental <MAT>

Query Match 21.7%; Score 876.5; DB 2; Length 603;  
Best Local Similarity 37.2%; Pred. No. 1.8e-54;  
Matches 202; Conservative 102; Mismatches 186; Indels 53; Gaps 15;  
Qy 111 SIDSIIVAASEODSEIMDANEQPAKVT--RSTVFTVTEAAPYAKSGGLDVCGLPIAL 168  
Db 47 SLNKLHVRTARATSGSDTSEKSLGKIVCGMSLVFVGAEVGPWSKTGGLDVLGLPPVL 106  
Qy 169 AARGHRVWVMPRYLNGSSDKNAKALYAKHIKIPCFGSGSHVTFHEYRDNDVAVFVD 228  
Db 107 AGNHRVMTVSPRY-----DQYKDAWDTNVLVEKVGDKIETVRFHCFKRGVDRFVD 160  
Qy 229 HPSY-----HRPGS-LYGDNFGA-FGDNQFRYTLCCYAAEAPLILELGG--YI---YQG 276  
Db 161 HPLFLERVWKTGSKLGPDKTIDYRDNLRFSLCQAALEAPRVNLNLSKVFSGPYGE 220  
Qy 277 NCMEVVDNHASLPVLLAAKYPYGVYRDSRSTLVTHNLAHQGVPEPASTYPDLGLPPEW 336  
Db 221 DVEIVANDWHISALPCYLKSMYKSRGLYKNAKVAFCFTHNAYQGRNAFSDFLSNLPDEF 280  
Qy 337 YGALEWVFPPEWARRHALD-----KGEAVNPLKGVAVTADRIVTVSOGYSWEVTTABG-G 389  
Db 281 RSSDFDI-----DGYNKPCGKKIKNMKAGILLESQVFTVSPHYAKELISGEDRG 330  
Qy 390 QGLNELLSSRRKSVLNGIVNGIDINDWNPPTDKCLPHHYSVDLDS-GKAKCAELQKELGL 448  
Db 331 VELDNII--RSTGIIGVINGMDNREWSQTDRIYDVHNETTVEAKPLKGLTQAEIGL 388  
Qy 449 PVREDVPLIGFGRGLDYQKGLDKLIMKALPELMREDVQFVMLGSDPIFGWMRSTESSYK 508

RESULT 14  
S11481  
UDPglucose--starch glucosyltransferase (EC 2.4.1.11) precursor - rice  
N:Alternate names: starch synthase; waxy protein  
C:Species: Oryza sativa (rice)  
C:Date: 05-Mar-1995 #sequence\_revision 10-Nov-1995 #text\_change 15-Oct-1999  
C:Accession: S11481; S22519; S30485; PC2190; JQ2224  
R:Wang, Z.Y.; Wu, Z.L.; Xing, Y.Y.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.  
Nucleic Acids Res. 18, 5898, 1990  
A:Title: Nucleotide sequence of rice waxy gene.  
A:Reference number: JQ0703; MUID:91016948  
A:Accession: S11481  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-609 <WAN>  
A:Cross-references: EMBL:X53694; NID:g577598; PIDN:CAA3732.1; PID:g577599  
R:Okagaki, R.J.  
Plant Mol. Biol. 19, 513-516, 1992  
A:Title: Nucleotide sequence of a long cDNA from the rice waxy gene.  
A:Reference number: S22519; MUID:92322986  
A:Accession: S22519  
A:Molecule type: mRNA  
A:Residues: 1-609 <OKA>  
A:Cross-references: EMBL:X62134; NID:g20402; PIDN:CAA44065.1; PID:g20403  
R:Hirano, H.Y.; Sano, Y.  
Plant Cell Physiol. 32, 989-997, 1991  
A:Title: Molecular characterization of the waxy locus of rice (Oryza sativa).  
A:Reference number: S30485  
A:Accession: S30485  
A:Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-609 <HIR>  
A:Cross-references: EMBL:X58228; NID:g20400; PIDN:CAA41186.1; PID:g20401  
A:Accession: PC2190  
A:Molecule type: protein  
A:Residues: 78-113 <HI2>  
A:Experimental source: leaf, cDNA POSLHC2120  
C:Comment: This protein is involved in amylose synthesis in the rice endosperm.  
C:Genetics:  
A:Gene: waxy  
A:Introns: 113/3; 140/3; 173/3; 203/3; 225/1; 258/3; 295/2; 376/3; 435/3; 499/3; 528/3;  
C:Function:  
A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing  
A:Pathway: starch biosynthesis  
C:Superfamily: starch synthase  
C:Keywords: amyloplast; chloroplast; glycogen/starch biosynthesis; glucosyltransferase;  
F:1-77/Domain: transit peptide (amyloplast) #status predicted <TNP>  
F:78-609/Product: glycogen (starch) synthase #status experimental <MAT>

Query Match 21.7%; Score 877; DB 2; Length 609;  
Best Local Similarity 38.5%; Pred. No. 1.7e-54;  
Matches 205; Conservative 89; Mismatches 173; Indels 66; Gaps 15;  
Qy 140 STVFTVTEAAPYAKSGGLDVCGLPIALAAARGHRVWVMPRYLNGSSDKNAKALYAK 199  
Db 84 NVVFGAEMAPWSKTGGLGVLGLPPAMAANGHRVWVISPRY-----DQYKDAWDTSV 137  
Qy 200 HTKIPCFGSGSHVTFHEYRDNDVAVFVDHPSY-----HRPG-SLYGDNFGA-FGDNQFR 252  
Db 138 VAEIKVADRYERVFHCYKRGVORFIDHPSFLEKVGWKTGEKIYGPDTGVYKDNQMR 197  
Qy 253 YTLCLYAAEAPLILELGGY-----IYQNCMFVVDNHASLPVLLAAKYPYGVYRDS 307  
Db 198 FSLCQAALEAPRLNANNPNFKTYGDDVVFVNCNDWHTGTPLASLYLNKNTQPNIGYRNA 257  
Qy 308 RSTLVHNLAHQGVPEASTYPDLGLPPEWYGALEWVFPPEWARRHALD--KGEAVNELKG 364  
Db 258 KVAFCHNISYQGRFAFEDYPELNLSEFRSSFDI-----DGYDTPVEGKINWMA 310  
Qy 365 AVVTADRIVTVSQGYSWE-VTTAEGGQGLNELLSSRRKSVLNGIVNGIDINDWNPPTDKCL 423







**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2002, 16:20:46 ; Search time 56.4 Seconds  
(without alignments)  
491.465 Million cell updates/sec

Title: US-09-674-824-2  
Perfect score: 4044  
Sequence: 1 MAATGVGAGCLAPSRLRAD.....SDGSLSVRTAEIRNLVTL 756

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 segs, 36664827 residues  
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3313	81.9	610	1 UGS2_WHEAT	Q43654 triticum ae
2	2609.5	64.5	626	1 UGS2_ORYSA	Q40739 oryza sativ
3	1992	49.3	641	1 UGS2_SOLTU	P93568 solanum tub
4	1204	29.8	752	1 UGS3_PEA	Q43093 pisum sativ
5	1154.5	28.5	788	1 UGS3_SOLTU	Q43847 solanum tub
6	932.5	23.1	608	1 UGST_MANES	Q43784 manihot esc
7	892.5	22.1	608	1 UGST_ANTWA	O82627 antirrhinum
8	884.5	21.9	607	1 UGST_SOLTU	O00775 solanum tub
9	883	21.8	608	1 UGST_SORBI	Q43134 sorghum bic
10	877	21.7	609	1 UGST_ORYSA	P19395 oryza sativ
11	876.5	21.7	603	1 UGST_PEA	Q43092 pisum sativ
12	872	21.6	609	1 UGST_ORYGL	Q42968 oryza glabe
13	870	21.5	605	1 UGST_MAIZE	P04713 zea mays (m
14	861	21.3	608	1 UGST_IPOBA	Q42857 ipomoea bat
15	858.5	21.2	603	1 UGST_HORVU	P09842 hordeum vul
16	843.5	20.9	615	1 UGST_WHEAT	P27736 triticum ae
17	768.5	19.0	485	1 GLGA_BACST	O08328 bacillus st
18	756	18.7	484	1 GLGA_BACSU	P39125 bacillus su
19	696	17.2	480	1 GLGA_AGRTU	P39670 agrobacteri
20	685.5	17.0	477	1 GLGA_ECOLI	P08323 escherichia
21	626.5	15.5	476	1 GLGA_HAEIN	P45179 haemophilus
22	592.5	14.7	477	1 GLGA_SYNY3	P74521 synechocyst
23	556	13.7	1230	1 UGS4_SOLTU	Q43846 solanum tub
24	346	8.6	521	1 GLGA_METJA	Q39001 methanococc
25	214.5	5.3	2410	1 MOKI_SCHPO	Q9usK8 schizosacch
26	193	4.8	1369	1 MOKB_SCHPO	Q9y704 schizosacch
27	177.5	4.4	2397	1 MOKB_SCHPO	Q99854 schizosacch
28	175.5	4.3	2358	1 MOKD_SCHPO	Q9y719 schizosacch
29	173	4.3	2352	1 MOKC_SCHPO	O9uul4 schizosacch
30	127	3.1	377	1 CFSU_BACSU	P46915 bacillus su
31	118.5	2.9	2205	1 POLN_RUBVT	P13889 rubella vir
32	118	2.9	351	1 LFSB_RHIME	Q9r9n2 rhizobium m
33	115	2.8	875	1 NETR_HUMAN	P56730 homo sapien

34	113.5	2.8	703	1	ULI7_HSV11	P10201 herpes simp
35	113.5	2.8	719	1	PBPA_STRPN	Q04707 streptococ
36	110.5	2.7	480	1	Y486_MYCTU	Q11152 mycobacteri
37	110.5	2.7	1045	1	SPS_BETVU	P49031 beta vulgar
38	110	2.7	484	1	P1GA_HUMAN	P37287 homo sapien
39	109.5	2.7	2142	1	BAT2_HUMAN	O48547 rhizobium l
40	109	2.7	352	1	LPCC_RHILV	O68547 rhizobium l
41	108	2.7	1490	1	CRK7_HUMAN	Q9yvv4 homo sapien
42	107.5	2.7	953	1	LYAG_MOUSE	P70699 mus musculu
43	107	2.6	493	1	Y911_CAUCR	P33976 caulobacter
44	107	2.6	589	1	RRFO_BPOBE	P14647 bacterioph
45	106	2.6	2715	1	TRX2_HUMAN	Q9umn6 homo sapien

ALIGNMENTS

```
RESULT 1
UCS2_WHEAT
ID UCS2_WHEAT STANDARD; PRT; 610 AA.
AC Q43654;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SOLUBLE GLYCOSYL [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11) (FRAGMENT).
OS Triticum aestivum (wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
OC Triticaceae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. T.A. FLORIDA; TISSUE=Endosperm;
RA Block N., Loez H., Lueticke S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] = UDP
CC + (1,4-ALPHA-D-GLUCOSYL)[N+1].
CC -!- PATHWAY: STARCH BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOSYL SYNTHASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U48227; BAB02197.1;
CC InterPro: IPR001296; Glycosyl_transf_1.
CC Pfam: PF00534; Glycosyl_transf_1;
CC GlycoGen biosynthesis; Transferase; Glycosyltransferase;
CC Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
CC NON_TER 1 1
CC TRANSIT <1 1 2 CHLOROPLAST (POTENTIAL).
CC CHAIN ? 610 SOLUBLE GLYCOSYL [STARCH] SYNTHASE.
CC BINDING 7 7 UDP-GLUCOSE (BY SIMILARITY).
CC SEQUENCE 610 AA; 67143 MW; 46080A3B7FB87193 CRC64;
CC -----
Query Match 81.9%; Score 3313; DB 1; Length 610;
Best Local Similarity 100.0%; Pred. No. 3.4e-227;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 147 EAAPYAKSGGLDVGCSLPIALAAARGHRVMMVMPRYLNGSSDKNYAKLYAKHIKIPCF 206
Db 1 EAAPYAKSGGLDVGCSLPIALAAARGHRVMMVMPRYLNGSSDKNYAKLYAKHIKIPCF 60
QY 207 GGSHEVTFEYEDRDNDVDFVDPSPYHRPGSLGVDNFGAFDGNQFRYTLCLCYAACAPLI 266
|||||
```



Db 61 GGSHEVTFHEYRDNDVDFVDPSPYHRRPGSLYDNGEAFGDNQFRYTLCLCYAAEAPLI 120  
QY 267 LELGGYIYQONCMFVYNDWHASLVPLVLLAAKRYGVRDRSSTLVHNLAHQGVPEAST 326  
Db 121 LELGGYIYQONCMFVYNDWHASLVPLVLLAAKRYGVRDRSSTLVHNLAHQGVPEAST 180  
QY 327 YPDGLGPPWYGALEWVFPWARRHALDKGEAVNFKGAVVTADRVTVSQYSWEVTTA 386  
Db 181 YPDGLGPPWYGALEWVFPWARRHALDKGEAVNFKGAVVTADRVTVSQYSWEVTTA 240  
QY 387 EGGQGLNELLSRKSRYLNGTVINGIDINDNPTTDKCLPHHYSVDDLSGKAKAEQKEL 446  
Db 241 EGGQGLNELLSRKSRYLNGTVINGIDINDNPTTDKCLPHHYSVDDLSGKAKAEQKEL 300  
QY 447 GLPVREDVPLIGRDLTVQKIDILKMAIPELMRDQVFMVGLSGDPIFEGWMRSTESS 506  
Db 301 GLPVREDVPLIGRDLTVQKIDILKMAIPELMRDQVFMVGLSGDPIFEGWMRSTESS 360  
QY 507 YDKFRGWGFSVPVSHRITAGCDILLMPSRFPCCGLNQLYAMQYGVTVPVVHGTGGLRDT 566  
Db 361 YDKFRGWGFSVPVSHRITAGCDILLMPSRFPCCGLNQLYAMQYGVTVPVVHGTGGLRDT 420  
QY 567 VETFPNPGKGEETGWAFSPLTVDKMLMALRTAMSTFREHKPSWGLMKRGMKTDHTWD 626  
Db 421 VETFPNPGKGEETGWAFSPLTVDKMLMALRTAMSTFREHKPSWGLMKRGMKTDHTWD 480  
QY 627 HAPSSSTRSSSGSWNTPTSCRGKGRKSCSPALKTSSTSSFRGPEGYPTLRCPATVE 686  
Db 481 HAPSSSTRSSSGSWNTPTSCRGKGRKSCSPALKTSSTSSFRGPEGYPTLRCPATVE 540  
QY 687 SQACCLWFAGSRITYDCAAAATVAGSGROLQFWGIRKKGCAAGMLTAKHHSDSLSVRVT 746  
Db 541 SQACCLWFAGSRITYDCAAAATVAGSGROLQFWGIRKKGCAAGMLTAKHHSDSLSVRVT 600  
QY 747 AEIRNQLVTL 756  
Db 601 AEIRNQLVTL 610  
RESULT 2  
ID UGS2\_ORYSA STANDARD: PRT: 626 AA.  
AC Q40739;  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE 15-JUL-1999 (Rel. 38, Last annotation update)  
DT SOLUBLE GLYCOPEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11) (SSS).  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 114-131.  
RC STRAIN=CV. JAPONICA; TISSUE=Seed;  
RX MEDLINE=94302151; PubMed=7518089;  
RA Baba T., Nishihara M., Mizuno K., Kawasaki T., Shimada H.,  
RA Kobayashi E., Ohnishi S., Tanaka K.-I., Arai Y.;  
RT "Identification, cDNA cloning, and gene expression of soluble starch  
synthase in rice (Oryza sativa L.) immature seeds.";  
RL Plant Physiol. 103:565-573(1993).  
CC -|- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] - UDP  
CC + (1,4-ALPHA-D-GLUCOSYL)[N+1].  
CC -|- PATHWAY: STARCH BIOSYNTHESIS.  
CC -|- SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE.  
CC -|- TISSUE SPECIFICITY: LEAVES AND IMMATURE SEEDS.  
CC -|- MISCELLANEOUS: THREE FORMS OF SOLUBLE STARCH SYNTHASE WERE  
CC PURIFIED: RSS1, RSS2 AND RSS3.  
CC -|- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOPEN SYNTHASE  
CC FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
CC EMBL: D16202; BAA03739.1; -.  
DR InterPro: IPR001296; Glycos\_transf\_1.  
KW Pfam: PF00534; Glycos\_transf\_1; 1.  
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;  
KW Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.  
FT TRANSIT 1 113 CHLOROPLAST (POTENTIAL).  
FT CHAIN 114 626 SOLUBLE GLYCOPEN [STARCH] SYNTHASE, RSS3.  
FT CHAIN 122 626 SOLUBLE GLYCOPEN [STARCH] SYNTHASE, RSS1.  
FT BINDING 147 147 UDP-GLUCOSE (BY SIMILARITY).  
SQ SEQUENCE 626 AA; 68451 MW; 03E4182507D26658 CRC64;  
Query Match 64.5%; Score 2609.5; DB 1; Length 626;  
Best Local Similarity 82.6%; Pred. No. 2.5e-177; Indels 31; Gaps 9;  
Matches 497; Conservative 32; Mismatches 42;  
QY 2 AATGVGAGCL-APSVRLRADPATAARASACVVRRLRLRLARGRYVAELSRGPAARPAQ 60  
Db 5 AGMGGAACLVAPQVR-----PGR-----RLRLQVRR-RCVAELSRDGGSA-----H 46  
QY 61 QQLAP-----PLVPGFLAP---PPAPAQS---PAPTQPLPDAGVGLAPDLLLEGIAED 110  
Db 47 GLPALAPLVKQVPLTFLVPTSTTPAPTQSPAPATPPPLPDGSGVEIEPD---LEGLED 104  
QY 111 SDISIIVAASEQDSEIMDANEQOAKVTRISIVFTGEAAPAKSGGLGDCGSLPATAAA 170  
Db 105 SDIKTIFVASEQSEIMDVKEQAQKAVTRSVFVFTGEASPYAKSGGLGDCGSLPATAAL 164  
QY 171 RGRHVVMVPRYLNGSSDKNYAKALYTAHKIKIPCGGSHEVTFEHEYRDNDVDFVDP 230  
Db 165 RGRHVVMVPRYMNALNKNFANAFVTEHKIKIPCGGSHEVTFEHEYRDNDVDFVDP 224  
QY 231 SYHRPGSLYDGNFGAGDNQFRYTLCLCYAAEAPLIELGGYIYQONCMFVYNDWHASLV 290  
Db 225 SYHRPGSLYDGNFGAGDNQFRYTLCLCYAAEAPLIELGGYIYQONCMFVYNDWHASLV 284  
QY 291 PVLLAAKRYPGYVYRDRSRTLVTHNLAHQGVPEASTYDGLGPPWYGALEWVFPWARR 350  
Db 285 PVLLAAKRYPGYVYRDRSRLVHNLAHQGVPEASTYDGLGPPWYGALEWVFPWARR 344  
QY 351 HALDKGEAVNFKGAVVTADRVTVSQYSWEVTTAEGGQGLNELLSRKSRYLNGTVINGI 410  
Db 345 HALDKGEAVNFKGAVVTADRVTVSQYSWEVTTAEGGQGLNELLSRKSRYLNGTVINGI 404  
QY 411 DINDNPTTDKCLPHHYSVDDLSGKAKAEQKELGLPVREDVPLIGRDLTVQKID 470  
Db 405 DINDNPTTDKCLPHHYSVDDLSGKAKAEQKELGLPVREDVPLIGRDLTVQKID 464  
QY 471 LKMAIPELMRDQVFMVGLSGDPIFEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCD 530  
Db 465 LKMAIPELMRDQVFMVGLSGDPIFEGWMRSTESSYKDKFRGWGFSVPVSHRITAGCD 524  
QY 531 ILLMPSRFPCCGLNQLYAMQYGVTVPVVHGTGGLRDTVETNPFGAKGEGCTGWAFSPLTV 590  
Db 525 ILLMPSRFPCCGLNQLYAMQYGVTVPVVHGTGGLRDTVETNPFGAKGEGCTGWAFSPLTV 584  
QY 591 DK 592  
Db 585 EK 586  
RESULT 3  
ID UGS2\_SOLTU STANDARD: PRT: 641 AA.  
AC P93568;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)



QY	102	LLLEGIAEDS--IDSIIIVASEQ--SEIMDANEQOQAKVYRSIVFTGAAAPYAKSGGL	157
Db	90	LiPHSVAGDATMVESHDIIVANDDLSEDETEEMEETPIKLTNIIIFVTAEAAPYSKTGGL	149
QY	158	GDVCGSLPIALAAHRHVMVMPRYLNGS--SDKNYAKALYAKHIKIPCFCGSGSHEVTFH	216
Db	150	GDVCGSLPMLAALAGRHRVMVSPRYLNGGSDEKIYANAVDLVRATVHCFGDGAQEAIFYH	209
QY	217	EYRONVDWFVDHPSYHRPGLSLGDNFGAFGDNQFRYTLTLLCYAACAPLLELGLGYTGO	276
Db	210	EYRAGVDWFVDHSSYCRPCTPYGDIYAGFDNQFRFTLLSHAACEAPLVPLGGFTYGE	269
QY	277	NCFVYVNDWHASLVVPLLAAKYRPYGVYRDSRSTLVIHNLAHOGVEPASTYDOLGLPPEW	336
Db	270	KCLFLANDHAAVLPLLLAAKYPYGVYKDAISVAIHNTAHOGVEPAVTYNLGLPPQW	329
QY	337	YGAEWVFPWARRHALDKCAEYNFLKGAVVTADRIVTVSQGSWEYVTAEAGSGGNELL	396
Db	330	YGAVIEWIFPTWARAHALDTGETVNVLKGA.TAVADRILTVSQGSWEITTPEGYVGLHELL	389
QY	397	SSRKSVLNGIVNGIDINDWNPTTKCLPHRYSVDDLSGKAKCAEQLQKELGLPVREDVPL	456
Db	390	SSRQSVLNGTNGIDVNDWNPSTDEHIAHSYINDLSGKVQCKTDLQKELGLPIRPDCPL	449
QY	457	IGFIRGLDYOKGIDLKMAIPELMREDVQVFMVLGSGDPIFEGWMRSTESSYKOKFRGWG	516
Db	450	IGFIRGLDYOKGVDIILSAIPELMNDQVQVFMVLGSGEKQYEDNMRTTENIFLCKOKFRAWG	509

Query Match	29.8%	Score 1204;	DB 1;	Length 752;
Best Local Similarity	47.9%;	Pred. No. 1.2e-77;		
Matches 256; Conservative 74;	Mismatches 169;	Indels 36;	Gaps 12;	
QY	105	EGTAEDSIDSIIVAAEQDSEIMDNFOPQAQVTR--SIVFVTGEAAPYAKSGGLGDVCG	162	
		:   : : :	:   :   :   :   :	
Db	228	EGANEPSSKEV--ANEAEFNESGGEKPPPLACTNVMNIILVSAECAPWSKTGGLGDVAG	284	
		:   : : :	:   :   :   :   :	
QY	163	SLPTALAAARGHRVAVMVPRLNGSSDKNAKA--LVTAKHIKIPCFGSHSEVFTFEHYRD	220	



```
Db 285 SLPKALARRGRHVMYVAPHY-----GNTAEAHIDIGVRKRYKVA--GQDMVEVTFYHTYID 336
Qy 221 NVQWVFDHPSYHR-PCSLYGDNFAGDNQFRYTLILCYAACAPLILELGGYIYQ-NC 278
Db 337 GVDIVFIDSPFNLESNIYGN--RDLILRRMWLFCAAVEVPHVPCGGICYGDGNL 393
Qy 279 MFVVDWHAASLVPLVLAARPYGVYRDSRSTLVIINLAHQGVPEASTYVDLGLPWEYWG 338
Db 394 VFIANDWHTALLPVYLKAYYRDHGLMNYTRSLVLIHIAHQGRGPEVDFNTVDLSGNYLD 453
Qy 339 ALQWVPEWARRHALDKGAENVLKGAVVTADRIYVTSOGYSWETVTAEGGGLNELLSS 398
Db 454 LFKWDPV-----GGEHNFIFAAGLTADRIYVSHGYAWELKTSEGWGLHNLNE 505
Qy 399 RKSVLNGIYVINGIDINWNPPTDKLPH-----HYSVDL-SGKAKCAEQLQKELGLPVRED 453
Db 506 SDMKFRGIVNGVDTKWNQFADLYTSDGYTNYNLKTLOTGRQCKAAALQRELGLPVRED 565
Qy 454 VPLIGFIRLDYOKGIDLKMAPELMDREDVQVLMGSDGPIPEGWMRSTESSYKDKFRG 513
Db 566 VPIISFIRLDYOKGIDLKMAPELMDREDVQVLMGSDGPIPEGWMRSTESSYKDKFRG 625
Qy 514 WGFSPVSHRITAGCDILLMPSPFPCGGLNOLYAMOYGVVPHVHGTGGRLDRVTFENPF 573
Db 626 WGFSPVSHRITAGCDILLMPSPFPCGGLNOLYAMOYGVVPHVHGTGGRLDRVTFENPF 685
Qy 574 GAKGECTGWAFSPLVTDKMLWALRTAMSTFREHKPSWEGMLKRGMTKDHDTWHA 628
Db 686 ---DESGVGWTFDRAEANKLMAALWCLLTYKDYKKSWEIGRGSQDLSWDNA 737

RESULT 5
UGS3_SOLUTU STANDARD; PRT; 788 AA.
ID UGS3_SOLUTU STANDARD; PRT; 608 AA.
AC Q43847;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCOPEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11) (GBSSI) (GRANULE-
DE BOUND STARCH SYNTHASE II) (FRAGMENT).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 66-75.
RC STRAIN=CV. DESIREE; TISSUE=Tuber;
RX MEDLINE=95400340; PubMed=7670507;
RA Edwards A., Marshall J., Sidebottom C., Visser R.G.F., Smith A.M.,
RA Martin C.;
RT "Biochemical and molecular characterization of a novel starch
RT synthase from potato tubers.";
RL Plant J. 8:283-294(1995).
CC -1- FUNCTION: ACCOUNTS FOR ONLY 10 TO 15% OF THE TOTAL SOLUBLE STARCH
CC SYNTHASE ACTIVITY IN TUBERS.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] =
CC UDP + (1,4-ALPHA-D-GLUCOSYL)[N+1].
CC -1- PATHWAY: STARCH BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST/AMYLOPLAST, SOLUBLE AND GRANULE-
CC BOUND.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOPEN SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X87988; CAA61241.1; .
```

```
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00334; Glycos_transf_1;
KW Glycogen biosynthesis; Transference; Glycosyltransferase;
KW Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
FT NON_TER 1 1 CHLOROPLAST.
FT TRANSIT <1 65 GLYCOPEN [STARCH] SYNTHASE.
FT CHAIN 66 788 UDP-GLUCOSE (BY SIMILARITY).
FT BINDING 310 310
FT VARIANT 71 71 S -> D.
SQ SEQUENCE 788 AA; 87890 MW; 8D8B90611B862B7B CRC64;

Query Match 28.5%; Score 1154.5; DB 1; Length 788;
Best Local Similarity 43.7%; Pred. No. 4.1e-74;
Matches 253; Conservative 83; Mismatches 194; Indels 49; Gaps 12;

Qy 79 PAPSAPTPTPL-----PDAGYGLAPDLLLLLEGIA-----EDSIDSIIVAAS 120
Db 215 PPTPSSQETLLDVSRSKSLVDVPGKKIQSYMPSLRKSSASHVQORNELEGSSAEAN 274
Qy 121 EQDSEIMDANEQPOAKV---TRSFVPTGEAAAPYAKSGGLGVDVCGSLPTALAAARGHRVW 177
Db 275 EETEDPVNIDKPPPLAGTNVMNIILVASECAPWSKTGGLGVAGALPKALARRRHRVW 334
Qy 178 VMPRYLNGSSDKNYAKALYAKHIKIPCFGGSHEVTFHFHEYRDNDVDFV-DHPSYHRPG 236
Db 335 VAPRYDNYPEPQDS---VRKTYKYD--GQDVDTYFQALLMDCDFVFIHSMFRIGN 388
Qy 237 SLYGDNFAGDNQFRYTLILCYAACAPLILELGGYIYQ-NCMFVVDWHAASLVPLVLA 295
Db 389 NIYGN---RVDILKRWLFCAAEVPHVHGTGGRLDRVTFENPF 445
Qy 296 AKRYGVYRDSRSTLVIINLAHQGVPEASTYVDLGLPPEWYGALEWVPEWARRHALDK 355
Db 446 AYYRNGIMNYTRSLVLIHIAHQGRGPLEDFSYVDLPHYMDPFLYDPV-----G 497
Qy 356 GEAVNELKGAIVTADRIYVTSOGYSWETVTAEGGGLNELLSSRKSVLNGIYVINGIDNW 415
Db 498 GEHNFIFAAGLTADRIYVSHGYAWELKTSEGWGLHNLNEQWKLQGLVINGIDITKEW 557
Qy 416 NPPTDKLPH-----HYSVDL-SGKAKCAEQLQKELGLPVREDVPLIGFIRLDYOKGI 469
Db 558 NPGLDVHLPRSDGYMNYSLDTLOTGRQCKAAALQKELGLPVREDVPLIGFIRLDYOKGI 617
Qy 470 DLKMAPELMDREDVQVLMGSDGPIPEGWMRSTESSYKDKFRGWGVSFVPSHRITAGC 529
Db 618 DLIAEAPVPMGQDVOLVLMGTGRDLQMLRQFECQNDKIRGWGVSFVPSHRITAGC 677
Qy 530 DILLMPSPFPCGGLNOLYAMOYGVVPHVHGTGGRLDRVTFENPFEGAKGEGTGWAFSPLT 589
Db 678 DILLMPSPFPCGGLNOLYAMOYGVVPHVHGTGGRLDRVTFENPFEGAKGEGTGWAFSPLT 734
Qy 590 VDKMLWALRTAMSTFREHKPSWEGMLKRGMTKDHDTWHA 628
Db 735 ASQILPRINCLLYREYKKSWEIGIOTRCMTQDLSWDNA 773

RESULT 6
UGS3_MANES STANDARD; PRT; 608 AA.
ID UGS3_MANES STANDARD; PRT; 608 AA.
AC Q43784;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GRANULE-BOUND GLYCOPEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).
GN WAXY OR GBSS.
OS Manihot esculenta (Cassava) (Manioc).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids I; Malpighiales; Euphorbiaceae; Manihot.
OX NCBI_TaxID=3983;
RN [1]
RP SEQUENCE FROM N.A.
```



```
QY      586 SPLTVD 591  
       | |  
Db     522 GALHVE 527
```

```
RESULT   7  
UGST_ANTMA          STANDARD;             PRT;    608 AA.  
AC O82627;  
AT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DI 15-JUL-1999 (Rel. 38, Last annotation update)  
DE GRANULE-BOUND GLYCAGEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).  
GN WAXI OR GBSS.  
OS Antirrhinum majus (Garden snapdragon).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OX Asteridae; euasterids I; Lamiales; Veronicaceae; Antirrhinum.  
OX NCBI_TaxId=4151;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Leaf;  
RA Merida A., Rodriguez J.M., Vincent C., Romero J.M.;  
RT "The granule-bound starch synthase (GBSS) gene from Antirrhinum majus  
is developmentally and circadian-clock regulated";  
RL Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.  
CC -!- CATALYTIC ACTIVITY: UDP-GLUCOSE + {1,4-ALPHA-D-GLucosyl}[N] = UDP  
        + ({1,4-ALPHA-D-Glucosyl}[N+1]).  
CC -!- PATHWAY: STARCH BIOSYNTHESIS.  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.  
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL PLANT GLYCEN SYNTHESE  
        FAMILY.
```

```
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial/  
entities requires a license agreement (See http://www.isb-sib.ch/announce/or  
or send an email to licensed@lsb-sib.ch).
```

```
EMBL; AJ006293; CAAG6958.1; -.  
DR EMBL; AJ006294; CAAG6959.1; -.  
InterPro; IPR001296; Glycos_transf_1.  
DR Pfam; PF00334; Glycos_transf_1; 1.  
KW Glycogen biosynthesis; transferase; Glycosyltransferase;  
FT Transit peptide; Chloroplast; Starch biosynthesis.  
TS TRANSIT           1 78 CLOROPLAST (BY SIMILARITY).  
FT CHAIN            79 608 GRNULE-BOUN GLYCEN [STACH] SYNTHASE.  
FT BINDING         95 96  UPD-GLUCOSE (BY SIMILARITY).  
SQ SEQUENCE        608 AA; 66361 MW; 6527D53D565B6E0C CRC64;
```

```
Query Match              22.1%; Score 892.5; DB 1; Length 608;  
Best Local Similarity 39.5%; Pred. No. le-55;  
Matches 204; Conservative 96; Mismatches 180; Indels 37; Gaps 15;
```

```
QY 130 NEPQAQT---RSIVFVTEGAAPYAKSGGLGDVCGLSPIALAARGHRVMYMRYNLNG 185  
| | : : ::||:: ||:::||||:: ||:: ||:: ||||:: ||:  
Db 69 NGSLPKGIIGCTGMNVLFVLAEVGPMWKTGGLDGVVGGLPAMAGNHGRVTMSPRY-- 125  
  
QY 186 SSDKNYKALYTAKH.I.KIPCFGSGSHVTFPEHYRDNDVNDFVFDPHSY-----HRPGSLY 239  
| | : | : | : ::|| | | | | | | | | | | | | : | | | | | : | |  
Db 126 ---DQRKDAMDQTVSVVVIEIKVKDGDSIETVRPFCHYKRGVDRVFRVFDHPIFLEKRWGKTSKIY 182  
  
QY 240 GDNFGE-FGDNQRYRTLCLCYAAACEAPLIHEL---GYEI---YGONCNFMVMDHWASLVPVLT 293  
| | : | : | : ||| ||| | | | | | | | | : | : | | | | | | : | |  
Db 183 GPNAGTDYODNQLRFSLSLCOALEAFRVNLNLTSSKYFGSYGEDVGFVANWHHTALLPCY 242  
  
QY 294 LAARYRGYVYDRSRSTLVIIHNLAHGCVPEASTYPDLGLPPWEWGVALEWWFPWARHAL 353  
| | : | : | : ||| ||| | | | | | | | | : | : | : | : | : | : |  
Db 243 LKSMYOGSKMYHLAKVAFCFIHNTAYQGRFGSSDFCLLNLPDQFKSSPDF--FDGYEKPY--- 298
```



```
QY 354 DKGEAVNFKGAVVADRTVTSQGSWE-VTTAEGCGQLNELLSSRKSXVNLGIVNGIDI 412
|| :||:| :||:||||| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 299 VKGRKINMKGAGLSDRVVTSPPYAMELVSGAERGVLDNNVIA--KTSITGIVNGMDT 356
|| :||:| :||:||||| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 413 NDNNPTTDCPLPHYSVDD--LSGKAKCKAELOKELGLPVREDVPLIGFGRIDYQKIDL 471
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 357 QENWPATDKHIDTNYDITVMDAKPLKALQANAGLPVDKNIPVIGFGRLEEQKGSIDI 416
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 472 IKMAIPELMREDVQFVNLGSGDP IFEGWMRSTESSYKDKFRGMVGFSPVSHRITAGCIDI 531
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 417 LVAAISKEFGLDVQIIILGTGKKKEEQIQELEVLPDKARGVAKFNVLPHAMITAGADF 476
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 532 LLMPSEFCGCLNQLAMQYCTVPVHGTGGLRDTV-ETFNPF--CAKEEGTGWAFSPL 588
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 477 MLVPSRFEPCQLQLHAMRYGTIPICASTGLVDTVTGTFGHMGAFNVECA--TVDPA 534
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 589 TVDKMLWALRTAMSTFREHKPSWEGMLKRGMTKDHWTW 625
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 535 DVOKIATTVERALAAY--GSVAIKEMIQNCMAQQLSW 569
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|

RESULT 8
UGST_SOLTU STANDARD; PRT; 607 AA.
AC Q00775; Q43176;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GRANULE-BOUND GLYCOPEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).
GN WAXY OR GBSS.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=AW79, 7322; PubMed=1896609;
RX MEDLINE=91360072; PubMed=1896609;
RA van der Leij F.R.; Visser R.G.F.; Ponstein A.S.; Jacobsen E.,
RA Feenstra W.J.;
RT "Sequence of the structural gene for granule-bound starch synthase of
RT potato (Solanum tuberosum L.) and evidence for a single point
RT deletion in the anf allele."
RL Mol. Gen. Genet. 228:240-248(1991).
RN [2]
RP REVISIONS.
RA van der Leij F.R.;
RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DONGNONG303;
RA Dai W.L.; Deng W.; Gai M.; Xiu M.; Zhao S.Y.; Wang X.M.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] = UDP
CC + (1,4-ALPHA-D-GLUCOSYL)[N+1].
CC -I- PATHWAY: STARCH BIOSYNTHESIS.
CC -I- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOPEN SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: X58453; CAA41359.1; -
CC DR EMBL: X83220; CAA56220.1; -
CC DR F01; S16555; YUPOY.
CC InterPro: IPR001296; Glycos_transf_1.
```

```
DR PF00534; Glycos_transf_1; 1.
KW Glycogen biosynthesis; transferase; Glycosyltransferase;
KW Transit peptide; Chloroplast; Starch biosynthesis.
FT TRANSIT 1 77
FT CHAIN 78 607
FT BINDING 95 95
FT CONFLICT 130 130
FT CONFLICT 338 338
FT CONFLICT 398 398
FT SEQUENCE 607 AA; 66575 MW; 2A377865CFAFA650 CRC64;
SQ
Query Match 21.9%; Score 884.5; DB 1; Length 607;
Best Local Similarity 39.1%; Pred. No. 3.9e-55;
Matches 202; Conservative 97; Mismatches 158; Indels 59; Gaps 15;
QY 140 SIVFVTGEAAPAKSGGLDGVCGSLPDLAARGHVVMVMPYVYNGSSDKYAKALYTAK 199
|| :||:| :||:||||| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 82 NLIFVTEVGPWSKTGGLDGLGGLPDLAARGHVMTISPRY-----DQYKADMTSV 135
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 200 HKIPCFGSGHEVTFPEHYRDNDVWVVDHPSY-----HRPGS-LYGDNFGA-FGDNQFR 252
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 136 AVEVKVGDSTIEIVRFHFCYKRGVDRVFDHPMFLEKVGWGTGSKYGPAGLDYLDNEIR 195
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 253 YTLICVAACEAPLILEL--GGYI---YGONCMFVNDWHASLPVLLAAKYRYPGVYRDS 307
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 196 FSLCQAALEAPKVLNLSNYSFGPYGEDVLFANDWHTALIPCYLKSWMYQSRGIYLN 255
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 308 RSTLVIIHNAHQGVPEASTYPDLGLPEWYGALEWYFPEWARRHALD-----KGBAVNF 361
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 256 KVAFCIHNIAYQGRSFSDFFLLNLPDEFGRSDFI-----DGYEKPVKGRKINW 305
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 362 LKAGAVTADRIVTSQGSYWEVTTA-EGGQGLNELLSRKSXVNLGIVNGIDINDWPTTD 420
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 306 MKAGIESHRVTVTPSYAQELVSAYDKGVELDSVL--RRTCTIGVINGMDTQEWNPATD 363
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 421 KCLPHHSYVDD--LSGKAKCKAELOKELGLPVREDVPLIGFGRIDYQKIDLKMAPEL 479
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 364 KYTDVKYDITVMDAKPLKALQAAVGLPVDKKIPILGIFGRLEQKSGDILVAAIHKF 423
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 480 MREDVQFVLMGSGDPIPEGWMRSTESSYKDKFRGMVGFSPVSHRITAGCDILLMPSRFE 539
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 424 IGLDVQIVLGTGKKEFEQIEQLEVLYPNKAQVAKFNVPPLAHMITAGADFLVPSRFE 483
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 540 PCGLNOLYAMOYCTVPVHGTGGLRDTVETFPNFGAKGEGTGW-----AFS-----PLT 589
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 484 PCGLIQHAMRYGTVPICASTGGLVDTVK-----EGYTGFFHMGAFNVECDVDPAD 534
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
QY 590 VDKMLWALRTAMSTFREHKPSWEGMLKRGMTKDHWTW 625
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
Db 535 VLKIVTTVARALAVY--GTLAFAEMIKNCMSELSW 568
|| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|

RESULT 9
UGST_SORBI STANDARD; PRT; 608 AA.
AC Q43134;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GRANULE-BOUND GLYCOPEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).
GN WAXY OR WX.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. 12311; TISSUE=Seed;
RA Hsing Y.C.; Liu C.; Yu H.; Hsieh J.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] =
CC UDP + (1,4-ALPHA-D-GLUCOSYL)[N+1].
```



01-NOV-1990 (Rel. 16, Created)  
01-NOV-1990 (Rel. 16, Last sequence update)  
15-JUL-1999 (Rel. 38, Last annotation update)  
GRANULE-BOUND GLYCAGEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).  
WAXY OR WX.  
Oryza sativa (Rice).  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
NCBI\_TaxID=4530;  
[1]  
SEQUENCE FROM N.A.  
RP TISSUE=Seed;  
RC MEDLINE=92322986; PubMed=1377969;  
RX Okagaki R.J.;  
RA "Nucleotide sequence of a long cDNA from the rice waxy gene.";  
RT Plant Mol. Biol. 19:513-516(1992).  
RL [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=CV. JAPONICA TAICHUNG 65; TISSUE=Seedling;  
RC Hirano H.Y., Sano Y.;  
RA "Molecular characterization of the waxy locus of rice (Oryza  
RT sativa).";  
RL Plant Cell Physiol. 32:989-997(1991).  
RN [3]  
RN SEQUENCE FROM N.A.  
RP STRAIN=CV. JAPONICA, AND CV. HANFENG;  
RC MEDLINE=91016948; PubMed=2216792;  
RX Wang Z.Y., Wu Z.L., Xing Y.Y., Zheng F.G., Guo X.L., Zhang W.G.,  
RA Hong M.M.;  
RT "Nucleotide sequence of rice waxy gene.";  
RL Nucleic Acids Res. 18:5898-5898(1990).  
RN [4]  
RN SEQUENCE FROM N.A.  
RC STRAIN=CV. INDICA;  
RA Wang X.Q., Wang Z.Y., Hong M.M.;  
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RN SEQUENCE OF 153-343 FROM N.A.  
RX MEDLINE=91200672; PubMed=2016064;  
RA Shimada H., Tada Y.;  
RT "Rapid isolation of a rice waxy sequence: a simple PCR method for the  
analysis of recombinant plasmids from intact Escherichia coli  
cells.";  
RL Gene 98:243-248(1991).  
CC -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.  
CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + {1,4-ALPHA-D-GLUCOSYL}[N] = UDP  
CC + {1,4-ALPHA-D-GLUCOSYL}[N+1].  
CC -1- PATHWAY: STARCH BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.  
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCAGEN SYNTHASE  
CC FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X62134; CAA44065.1; -  
DR EMBL; X58228; CAA41186.1; -  
DR EMBL; X53694; CAA37732.1; -  
DR EMBL; X65183; CAA46294.1; -  
DR EMBL; M55039; AAA33918.1; -  
DR PIR; JQ0703; JQ0703.  
DR PIR; S22519; S22519.  
DR PIR; S30485; S30485.  
DR PIR; JQ2224; JQ2224.  
DR InterPro; IPR001296; Glycos\_transf\_1.  
DR Pfam; PF00534; Glycos\_transf\_1; 1.  
DR GlycoGen biosynthesis; Transferase; Glycosyltransferase.



```

KW  Transit peptide; Chloroplast; Starch biosynthesis.
FT  CHAIN 1 77 CHLOROPLAST.
FT  BINDING 78 609 GRANULE-BOUND GLYCOCEN [STARCH] SYNTHASE.
FT  CHAIN 97 97 UDP-GLUCOSE (BY SIMILARITY).
FT  CONFLICT 247 247 N > T (IN REF. 5).
FT  CONFLICT 250 250 P > T (IN REF. 5).
FT  CONFLICT 415 415 P > S (IN REF. 4).
SQ  SEQUENCE 609 AA; 66476 MW; C225DBF6F12072C5 CRC64;

Query Match 21.7%; Score 877; DB 1; Length 609;
Best Local Similarity 38.5%; Pred. No. 1.3e-54;
Matches 205; Conservative 89; Mismatches 173; Indels 66; Gaps 15;

QY 140 SYFVPTGEAAPYAKSGGLGDCVCSLPLAALAGHRVVMVMPRYLNGSSDKNYAKALYAK 199
DB 84 NVFVGAEMAPWSKTGLGVLGGLPAMAGHRVVMISPRY-----DQYKDAWDTSV 137
QY 200 HIKTPCFGSHVETFFHEYRONVDVDFVDPHSY-----HRPG-SLYGDNFGA-FGDNQFR 252
DB 138 VAEIKVADRYVRFFHCYARGVDVDFHDPSPLEKVGKGTGEKIYGPDTGVQYKDNQMR 197
QY 253 YTLCYAACAPLILEGGY-----IYQNCMFVNDWHASLVPLLAARYRPGVYRDS 307
DB 198 FSLLCQAALAPRILNLLNNPYKGTGYGEDVVFVFCNDWHTGPLASYLKNYQPNGIYRNA 257
QY 308 RSLVLIHNLAHQGVPASTYDGLPLPENTYGALEWVPEWARRHALD---KGEAVNFKG 364
DB 258 KVAFCHINISYQGFAPFELNLSRFRSSFDI-----DGYDTPVEGRKINMKA 310
QY 365 AVTADRIVTVSGYSWE-VTTAEGGGLNELLSSRLVNGVINGIDINDWPTDKCL 423
DB 311 GILEADRLVSPYAEELISGTARGELDNIM--RUTGITGVINGMDVSEWDPKDYI 368
QY 424 PHRY-SVDDLGGKAKKAEQKELGLPVRDYPVIGFIRGLDYQKGLDLIKMAIPELMRE 482
DB 369 TAKYDATTAEAKALNEALQAEAGLPVDRKIPLIAFIAGRLEEQKGPVMAAIPELMQE 428
QY 483 DVQFVMLGSDPFEFGHMRSTESYKDKFRGWGVSFVPSVSHRTAGCDIILLMSRFEPCG 542
DB 429 DVQIVLGLTGKFKELKLMEEKYPGKVRVVKFNAPLAHLIMAGADVLAIVPSRFEPCG 488
QY 543 LNOLYAMOYGTVPVHGTGLRDTVETFNFPFGAKGEGTGWAFSPLEVD-----K 592
DB 489 LIQLOHRYGTCPACASTGGLVDTVI-----EGKTFHNGRSLVSDCKVVEPSDVKK 539
QY 593 MLWALRTAMSTFREHKPSWEGLMKRGMKGTHTWDHAPSSTSRSSGP--SWTN 643
DB 540 VAATLKRKRVV--GTPAYEEMVRNMCNQDLSW-----KCPAKKNEN 579

RESULT 11
UGST_PEA
ID UGST_PEA STANDARD; PRT; 603 AA.
AC Q43092;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GRANULE-BOUND GLYCOCEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11)
DE (GBSSI).
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3889;
RN [1]
RP SEQUENCE FROM N. A., AND SEQUENCE OF 76-88.
RC STRAIN=CV, BCI/RR; TISSUE=Embryo;
RX MEDLINE=93251108; PubMed=1302049;
RA Dry I., Smith A., Edwards A., Bhattacharyya B., Dunn P., Martin C.;
RT "Characterization of cDNAs encoding two isoforms of granule-bound
RT starch synthase which show differential expression in developing
RT storage organs of pea and potato.";
```

```

RL Plant J. 2:193-202(1992).
CC -1- FUNCTION: MAY BE RESPONSIBLE FOR THE SYNTHESIS OF AMYLOSE.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + {1,4-ALPHA-D-GLUCOSYL}[N] -
CC UDP + {1,4-ALPHA-D-GLUCOSYL}[N+1].
CC -1- PATHWAY: STARCH BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN PODS AND LEAVES. NO EXPRESSION
CC IN FLOWERS OR STIPULES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT ALL STAGES OF EMBRYONIC
CC DEVELOPMENT WITH HIGHEST LEVELS IN LATER DEVELOPMENTAL STAGES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOCEN SYNTHASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X88789; CAA61268.1; -.
CC InterPro: IPR001296; Glycos.transf_1.
CC Pfam: PF00534; Glycos.transf_1.1.
CC Glycogen biosynthesis; Transferase; Glycosyltransferase;
CC Transit peptide; Chloroplast; Starch biosynthesis.
CC TRANSIT 1 75 CHLOROPLAST.
CC CHAIN 76 603 GRANULE-BOUND GLYCOCEN [STARCH] SYNTHASE.
CC BINDING 91 91 UDP-GLUCOSE (BY SIMILARITY).
CC SEQUENCE 603 AA; 66362 MW; 817252FDD12CCAA0 CRC64;

Query Match 21.7%; Score 876.5; DB 1; Length 603;
Best Local Similarity 37.2%; Pred. No. 1.4e-54;
Matches 202; Conservative 102; Mismatches 186; Indels 53; Gaps 15;

QY 111 SIDSIIVAASEQSEIMDANEQPOAKVT--RSIVFVTGEAAPYAKSGGLGDCVCSLPLA 168
DB 47 SLNKLHVRTARATSGSSDTSSEKSLGKIVCGMSLVFVGAEGVPSWKTGLGDLGGLPVL 106
QY 169 AARGHRVVMVMPRYLNGSSDKNYAKALYAKHFKIPCGSGSHVETFFHEYRONVDVDFV 228
DB 107 AGNHRVMTVSPRY-----DQYKDAWDTNVLVEKVGDKLIEVRFHCYKRGVDRVFD 160
QY 229 HPSY-----HRPGS-LYGDNFGA-FGDNQFRYTLCLYAAEAPLILEGG--YI---YQ 276
DB 161 HPLFLERVWGKTGSKLYGPKTGIDYRDNLRLFSLLCQAALAPRVLNLSKYFSGPYGE 220
QY 277 NCMFVNDWHASLVPLLAARYRPGVYRDSRSTLVLIHNLAHQGVPEASTYDGLLPEW 336
DB 221 DVIFVANDWHASLIPCYLKSMSYKSLGKYNKAKVAFCTHNIAYQGRNAFSDFLNLPDEF 280
QY 337 YGALEWVPEWARRHALD-----KGEAVNFKAGVVTADRIVTVSGYSWEVTTAG-G 389
DB 281 RSSFDPI-----DGNKCEGKINMKAGILSDQVFTVSPHYAKELISGEDRG 330
QY 390 QGLNELLSRKSVLNGIVNGIDINDMPTDKLPHHYSVDLLS-GKAKCAELQKELGL 448
DB 331 VELDNII--RSTGLIGVINGMDNREWSPTDRVIDVHYNETTVEAKPLKGLTQAEIGL 388
QY 449 PVREDVPLIGFIRGLDYQKGLDLIKMAIPELMREDVQFVMLGSDPFEFGHMRSTESYK 508
DB 389 PVDSSPLIGFIRGLEQKSDILVEAIAKFAFADENQIVVLGTGCKIMEKIOIEVLEEKY 448
QY 509 DKFRGWGVSFVPSVSHRTITAGCDIILLMSRFEPCGLNQLYAMOYGTVPVHGTGLRDTVE 568
DB 449 GKAIGITKFNPSLAHKIIAGADFIIVPSRFEPCGLVQLHAMPYGTVPDVNSTGGLVDTVK 508
QY 569 T-----FNPFGAKGEGTGWAFSPLETVDKMLWALRTAMSTFREHKPSWEGLMKRGMK 622
DB 509 EGYTGFGHAGPDEFDECE-----VDPDDVDKLAATVVKRAKLYGTGTQ--AMKQIILNCMAQN 561
QY 623 HTW 625
```



```

RX MEDLINE=92134825; Pubmed=1685658;
RA Umeda M., Ohtsubo H., Ohtsubo E.;
RT "Diversification of the rice waxy gene by insertion of mobile DNA
RL elements into introns.";
RT Jpn. J. Genet. 66:569-586(1991).
CC -|- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.
CC -|- CATALYTIC ACTIVITY: UDP-GLUCOSE + {1,4-BETA-D-GLUCOSYL}[N] = UDP +
CC (1,4-BETA-D-GLUCOSYL)[N+1].
CC -|- PATHWAY: STARCH BIOSYNTHESIS..
CC -|- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -|- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCAGEN SYNTHASE
CC FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ or
CC send an email to licensel@isb-sib.ch).
CC -----
DR EMBL: D10472; BAA01272.1; -.
DR InterPro: IPR001296; Glycos_transf_1.
DR Pfam: PF00534; Glycos_Transf_1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
FT TRANSIT peptide; Chloroplast; Starch biosynthesis.
FT TRANSIT 1 77 CHLOROPLAST {BY SIMILARITY}.
FT CHAIN 78 609 GRANULE-BOUND GLYCAGEN [STARCH] SYNTHASE.
FT FT BINDING 97 97 UDP-GLUCOSE {BY SIMILARITY}.
SQ SEQUENCE 609 AA; 66475 MW; C228BFB9C407FA5 CRC64;
Query Match 21.6%; Score 872; DB 1; Length 609;
Best Local Similarity 38.3%; Pred. No. 3e-54;
Matches 204; Conservative 88; Mismatches 175; Indels 66; Gaps 15;

Qy 140 SIVFVTGEAAPAKSGGIGDVCGSLPTALAAHRVWVMRYLNGSSDKNYKALYTAK 199
   :||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 84 NVTFVGAEAPWSKTGGLGDVLGCLPMAANGRHVRVMSIPRY-----DOYKDWDTSV 137
   : ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Qy 200 HKIPCFFGGSHVTFHFHEYRDNDVMDVFVDHPSY-----HRPG-SLYGDNFGA-FGDNQFR 252
   : ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 138 VAEIKVADRYERVRFHCYKRGVDRVFDHPSPLEKVMGTGEIYGPDGTGYDKDNQMR 197
   : ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Qy 253 YTLACYACEAPLTILEGGY-----YQGNCMFVVNDWHASLVPLLAARYPPGVYRDS 307
   : ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 198 FSLQCQAALCAPRIILNNPNPFKTYTGDDVVFVCNDWHTGPLASYLNKNYPNGIYRNA 257
   : ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Qy 308 RSTLVIHNLAHOGVEPASTVPDGLPPEWYGALEGMVPEWARRALHD---KGEAVNFLKG 364
   : ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 258 KVAFCIIHSIQGFAREDPEDPELNSBRFSDFDI-----DCYDTPVPEGKINMKKA 310
   : ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |

```



```

RESULT 14
UGST_IPOBA
ID      UGST_IPOBA      STANDARD;          PRT;      608 AA.
AC 042857;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GRANULE-BOUND GLYCAGEN [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).
GN WAXY OR SS67.
OS Ipomoea batatas (Sweet potato) (Batatae).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Convolvulaceae; Ipomoea.
OX NCBI_TaxId=4120;
RN [1]
RP
RN
SEQUENCE FROM N.A.
RA STRAIN=CV. TAINONG 57; TISSUE=Tuberous root;
RA Wang S.J., Yeh K.W., Tsai C.Y.;
RL Submitted (DEC-1995) to the EMBL/Genbank/DBD databases.
CC -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] = UDP
CC -1- + (1,4-ALPHA-D-GLUCOSYL)[N+1].
CC -1- PATHWAY: STARCH BIOSYNTHESIS.

```

```

CC -1- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOSYL SYNTHASE
CC FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U44126; AAA86423.1; -.
CC DR InterPro: IPR001296; Glycos_transf_1.
CC DR Pfam: PF00534; Glycos_transf_1.1.
CC KW Glycosyl biosynthesis; Transferase; Glycosyltransferase;
CC KW Transit peptide; Chloroplast; Starch biosynthesis.
CC FT TRANSIT 1 76 CHLOROPLAST (BY SIMILARITY).
CC FT CHAIN 77 608 GRANULE-BOUND GLYCOSYL [STARCH] SYNTHASE.
CC FT BINDING 96 95 UDP-GLUCOSE (BY SIMILARITY).
CC FT SEQUENCE 608 AA; 67000 MW; CCA8FA50A2F69CB0 CRC64;
CC -----
CC 'Query Match 21.3%; Score 861; DB 1; Length 608;
CC Best Local Similarity 38.9%; Pred. No. 1.8e-53;
CC Matches 210; Conservative 77; Mismatches 181; Indels 72; Gaps 17;
CC -----
QY 140 SIVFVTEAPYAKSGGLGVDVCGSLPTALAAGRHVWVMPRYLNGSSDKNYKALYTAK 199
   :||| | :|:||||| | ||||| ||||| ||||| | | | | | | | | | | | | | | | |
Db 83 NLVPGVCEEGPWCKTGGDLVGLGPPALAAARGHRVMTVCPY-----DQYKDAWETCV 136
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 200 HIKPCFGGSHE-VTFHEYRDNDVWVDFHPVS-----HRPGS-LYGDNFGA-FGDNQF 251
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 137 VVE-PQVGDRIEPRVFHSYKRGVDVFDHPHLEKVGWGTGSMVLGPKACKDYKDQL 195
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 252 RYTLACYAAEAPLILEGG--YI---YGNCMFVNDWHASLPVLLAAKRYPGVYRD 306
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 196 RFSLLCQAALAPRVLLNLSKYSFGYGEDVFVNDWHHTALLPCYLKTYQSRGIYMN 255
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 307 SRSTLVTHNLAHOGVEPASTYDGLGLEPWYGALEWVPEWARRHALD---KGEAVNFK 363
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 256 AKVAFCHNATYQGRFAFDFSLNLDEYKGSDFI-----DCYDKPVKGRKINMK 308
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 364 GAVYTADRIYTVSGYSWE-VTTAEGQGLNELSSRKSVLNGVINGIDINDWNPPTDKC 422
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 309 AGIREADRVFTSPYAKELVSCVSKGVELDNHI--RDCGITGICNGMDTQEWNPATDKY 366
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 423 LPHIYSVDD--LSGRKAKCAELQKELGVPREDVPLIGFGRDLYQKGLIDLKMAPELMR 481
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 367 LAVKYDITTVYQAARPLKEALQAQVLPVDRNRIPLIGFGRLEEQKSDILYAASKFIS 426
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 482 EDVGFVLMGSDPIFFGCMRSTESSYKDKFRGWGVFSVPVSHRTAGDILLMPSRPEPC 541
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 427 MDVQILILGTGKKFEQOIEOLEVWYDPAKRGVAKFNVPLAHMITAGADFMILISREPC 486
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 542 GLNQLYAMQYGTVPVWVHGTGGLRDTVTFNPFQKAGEEGTGWAFSPLTVD----- 591
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 487 GLIQLHAMRYGTPCICASTGGLVDTVK-----EGYTGFGHGAFNVDCEYDVPEDVL 537
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 592 KMLWALRATSTF-----REHKPSWEGLMKRGMTKDHDTWHDAPSTSRSSSGP 639
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 538 KVTITVGRALAIYGTLAFTEMIKMKCMQSELSWKKCPAK-----NWETVLLSLGVAGSEP 590
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
CC -----
CC RESULT 15
CC UGST_HORVU STANDARD; PRT; 603 AA.
CC ID UGST_HORVU AC P09842;
CC AC P09842;
CC DT 01-MAR-1989 (Rel. 10, Created)
CC DT 01-MAR-1989 (Rel. 10, Last sequence update)
CC DT 15-JUL-1999 (Rel. 38, Last annotation update)
CC DE GRANULE-BOUND GLYCOSYL [STARCH] SYNTHASE PRECURSOR (EC 2.4.1.11).
CC GN WAXY.

```



OS Hordeum vulgare (Barley).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;  
OC Triticeae; Hordeum.  
OX NCBI\_TaxID=4513;  
RN [1].  
RC SEQUENCE FROM N.A.  
RP STRAIN=CV. VOGELSANGER GOLD; TISSUE=Leaf;  
RX MEDLINE=88303345; PubMed=2970062;  
RA Rhode W., Becker D., Salami F.;  
RT "Structural analysis of the waxy locus from Hordeum vulgare.";  
RN Nucleic Acids Res. 16:7185-7186(1988).  
RN [2].  
RP SEQUENCE OF 76-89.  
RC STRAIN=CV. H354-295-2-5; TISSUE=starchy endosperm;  
RX MEDLINE=94170739; PubMed=8125056;  
RA Flengstad R.;  
RT "Separation of acidic barley endosperm proteins by two-dimensional electrophoresis";  
RL Electrophoresis 14:1060-1066(1993).  
CC -!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.  
CC -!- CATALYTIC ACTIVITY: UDP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)[N] -> UDP + (1,4-ALPHA-D-GLUCOSYL)[N+1].  
CC -!- PATHWAY: STARCH BIOSYNTHESIS.  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.  
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X07931; CAA30755.1; -.  
DR EMBL; X07932; CAA30756.1; -.  
DR PIR; S01727; YUBHY.  
DR InterPro; IPR001296; Glycos\_transf\_1.  
DR Pfam; PF00534; Glycos\_transf\_1.  
KW Glycogen biosynthesis; transferase; Glycosyltransferase;  
KW Transit peptide; Chloroplast; Starch biosynthesis.  
FT TRANSIT 1 72 CHLOROPLAST (BY SIMILARITY).  
FT CHAIN 73 603 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.  
FT BINDING 90 90 UDP-GLUCOSE (BY SIMILARITY).  
FT SEQUENCE 603 AA; 66211 MW; 0B0B3DE6A8217934 CRC64;  
SQ  
  
Query Match 21.2%; Score 858.5; DB 1; Length 603;  
Best Local Similarity 36.0%; Pred. No. 2.7e-53;  
Matches 214; Conservative 96; Mismatches 192; Indels 93; Gaps 18;  
  
QY 69 PGF--LAPPPAPA-----QSPAPTQPPIDAGVGEIAPDLLLEGIAEDSIDSIIVA 118  
DB 25 PCFQGLRPNRPADALGMRITGASAAAPKQSKAHGRSRCL-----SVVVS 70  
  
QY 119 ASEODSEINDANEQPOAKVTRSTVFVTEGAPYAKSGGLDVGSLPIALAARGHEVMV 178  
DB 71 ATGSG-----MNLVFGAEMAPWSKGTGGLDVLGGLPPAPAAANGHRVMV 115  
  
QY 179 MPRLNGSSDKNYAKALYAKHIKIPCFGSGHEVTFEYERONVDMVFVDHPISY----- 232  
DB 116 SPRY-----DQYKDAWDTSVISEIKVADEYERVFFHCYKRGVDRVFDHPWLEKVRG 169  
  
QY 233 HRPGLYGNFGA-FGDNQFRYTLICYAACEAPLILELGG--YI-----YQNCMFVYVNDWH 286  
DB 170 KTKETIYGDAGTDYEDNQRFSLCQAALAPRIILNANNPYFSGPYGEDVVFVVCNDWH 229  
  
QY 287 ASLVPVLLAAKVRPGYVGRSRLVIHNLAHQGVPEASTYDGLGLPPWYGALEWVPE 346  
DB 230 TGLLACYLKSQNSGIYRTAKVAFCIHNIYSQGRSFDDFAQLNLPFRKFSDFI--- 286

QY 347 WARRHALDK---GEAVNFKLGAVVADRTIVTVSQGYSWEVTTAEG-GQGLNELLSSRKSV 402  
DB 287 ----DCYDRPVEGRKINWKKAGILQADKVLTVSPYAEELISGEARGCCELDNIM--RLTG 340  
QY 403 LAGIVNGIDINDWNPPTDKCLPHYSVDD-LSGKAKCAKAELOKELGLPVREDVPLIGFIG 461  
DB 341 ITGIVNGMDVSEWDPTKDKFLAVNYDITTALEAKALNKEALQAEVGLPVDRKVPVAVFIG 400  
QY 462 RLDYOKGIDLIKMAIPELMR-EDVOFVMLGSGDPIFEGWNRSTESSYKDKFRGWVGSVP 520  
DB 401 RLEEQKGPDMVMTAAIPEILKEEDVQIILGTGKKKEFKLLKSMEEKFPQKVRVAVRFPNAP 460  
QY 521 VSHRITAGCDIILLMPSRFPCGLNQLYAMQYGTVPVYVHGTGGGLRDTVETFPFGAKGEG 580  
DB 461 LAHQMMAGADLLAVTSRFEPGGLIQLQGMRYGTPCVCASTGGGLVDIV-----EGK 511  
QY 581 TGWAFSPLTVD-----KMLWALRTAMSTFRHKPSWEGLMKRGMTKDHWTW 625  
DB 512 TGFHMGRLSVDGCVNVEPADVKVATTLKRAVKV--GTPAYQBMVKMNCMIQDLSW 564

Search completed: March 28, 2002, 16:30:31  
Job time: 585 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2002, 16:19:31 ; Search time 100.47 seconds  
(without alignments)  
1100.646 Million cell updates/sec

Title: US-09-674-824-2  
Perfect score: 4044  
Sequence: 1 MAATGVGAGCLAPSVRLRAD.....SDGSLSVRTAEIRNLVTL 756

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues  
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL17.\*

- 1: sp-archaea.\*
- 2: sp-bacteria.\*
- 3: sp-fungi.\*
- 4: sp-human.\*
- 5: sp\_invertebrate.\*
- 6: sp-mammal.\*
- 7: sp-mhc.\*
- 8: sp-organelle.\*
- 9: sp-phage.\*
- 10: sp-plant.\*
- 11: sp-rodent.\*
- 12: sp-virus.\*
- 13: sp-vertebrate.\*
- 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3355	83.0	647	10 Q9SOH0	Q9sqh0 aegilops ta
2	3351	82.9	647	10 Q9LECO	Q9leco triticum ae
3	3348	82.8	647	10 Q9SQG9	Q9sqg9 triticum ae
4	3343	82.7	647	10 Q9LEB9	Q9leeb9 triticum ae
5	3264	80.7	623	10 Q9MSA3	Q9msa3 hordeum vul
6	2617.5	64.7	623	10 Q9SSV9	Q9ssv9 oryza sativ
7	2617.5	64.7	626	10 Q24206	Q24206 oryza sativ
8	2591	64.1	622	10 Q49064	Q49064 zea mays (m
9	2562.5	63.4	629	10 Q9SBL2	Q9sbl2 sorghum bic
10	2038	50.4	652	10 Q9FNF2	Q9fnf2 arabidopsis
11	2025.5	50.1	575	10 Q9SEI7	Q9sei7 arabidopsis
12	1258	31.1	792	10 Q9MAC8	Q9mac8 arabidopsis
13	1245	30.8	630	10 Q65365	Q65365 ipomoea bat
14	1227.5	30.4	732	10 Q48899	Q48899 zea mays (m
15	1205.5	29.8	698	10 Q48900	Q48900 zea mays (m
16	1187	29.4	751	10 Q9SPH5	Q9sph5 manihot esc
17	1183.5	29.3	576	10 Q64926	Q64926 chlamydomon
18	1178.5	29.1	799	10 Q9LEE3	Q9lee3 triticum ae
19	1174.5	29.0	799	10 Q9SPW9	Q9spm9 triticum ae

20	1165	28.8	798	10 Q9LEE2	Q9lee2 triticum ae
21	1148.5	28.4	799	10 Q9M466	Q9m466 triticum ae
22	1128.5	27.9	491	10 Q24398	Q24398 triticum ae
23	930.5	23.0	606	10 Q9XIS6	Q9xis6 phaseolus v
24	890	22.0	708	10 Q64925	Q64925 chlamydomon
25	887.5	21.9	604	10 Q9PR03	Q9pr03 perilla fru
26	884	21.9	607	10 Q9ZSQ5	Q9zsq5 astragalus
27	868.5	21.5	604	10 Q9SLS6	Q9sls6 triticum tu
28	866.5	21.4	574	10 Q9SYU0	Q9syu0 triticum ae
29	863	21.3	610	10 Q9MAQ0	Q9maq0 arabidopsis
30	861	21.3	605	10 Q9SXX3	Q9sxx3 triticum ae
31	861	21.3	605	10 Q9FUD6	Q9fuu6 triticum ae
32	860	21.3	604	10 Q9SLS9	Q9sls9 triticum tu
33	859	21.2	604	10 Q9SLS7	Q9sls7 triticum ae
34	857	21.2	604	10 Q9S7N5	Q9s7n5 triticum ae
35	854.5	21.1	606	10 Q43012	Q43012 oryza sativ
36	852.5	21.1	605	10 Q9SQ51	Q9sq51 aegilops sp
37	848	21.0	605	10 Q9SQ52	Q9sq52 triticum mo
38	847	20.9	605	10 Q9SLS8	Q9sls8 triticum tu
39	844	20.9	604	10 Q9SXK4	Q9sxk4 triticum ae
40	826	20.4	599	10 Q9SQS8	Q9sq58 triticum ae
41	762.5	18.9	483	2 Q9CHM9	Q9chm9 lactococcus
42	760	18.8	565	10 Q9XEN9	Q9xen9 triticum ae
43	703.5	17.4	486	2 Q9WZZ7	Q9wzz7 thermotoga
44	654.5	16.2	476	2 Q9KDX6	Q9kdx6 bacillus ha
45	649.5	16.1	651	10 Q64927	Q64927 chlamydomon

## ALIGNMENTS

RESULT 1  
Q9SOH0 PRELIMINARY; PRT: 647 AA.  
ID Q9SOH0  
AC Q9SOH0  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE STARCH SYNTHASE I.

OS Aegilops tauschii (Tatropyzum tauschii).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;  
OC Triticeae; Aegilops.  
OX NCBI\_TaxID=37682;

RP SEQUENCE FROM N.A.  
RA Li Z., Rahman S., Kosar-Hashemi B., Mouille G., Appels R.,  
RA Morell M.K.;  
RT "Cloning and characterization of a gene encoding wheat starch synthase  
I.";  
RL Theor. Appl. Genet. 98:1208-1216(1999).

DR EMBL: AF091802; AAF03557.1;  
DR InterPro; IPR000531; TonB\_boxC.  
DR InterPro; IPR001296; Glycos\_transf\_1.  
DR Pfam; PF00534; Glycos\_transf\_1; 1.  
DR PROSITE; PS00430; TONB\_DEPENDENT\_REC\_1; UNKNOWN.1.  
SQ SEQUENCE 647 AA; 71018 MW; 9D4F791F8CE32728 CRC64;

Query Match 83.0%; Score 3355; DB 10; Length 647;  
Best Local Similarity 100.0%; Pred. No. 9.2e-241;  
Matches 628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAATGVGAGCLAPSVRLRADPATARASACVVRRLRLRLRGLRGRVVAELSRGPAARPAQ 60  
Db 1 MAATGVGAGCLAPSVRLRADPATARASACVVRRLRLRLRGLRGRVVAELSRGPAARPAQ 60  
Qy 61 QQLAPPLVPGFLAPPPAPAPQAPPTOPPLPDAGVGLAPDLLLEGIAESIDSIIVAAS 120  
Db 61 QQLAPPLVPGFLAPPPAPAPQAPPTOPPLPDAGVGLAPDLLLEGIAESIDSIIVAAS 120  
Qy 121 EQDSEIDANDNEQPAKYTRSVIVFTGGAAPYAKSGGLGCVGSLPIALAAARGHRVMVMP 180  
Db 121 EQDSEIDANDNEQPAKYTRSVIVFTGGAAPYAKSGGLGCVGSLPIALAAARGHRVMVMP 180



```
Db 121 EQDSEIMDANEQPAQKVTSTVFTGEAAPYAKSGGLGVDGSLPDLAALAAHRVVMVMP 180
Qy 181 RYLGSSDKNYAKALYAKHIKIPCGGSHVETFFHEYRDNDVDFVDPHPSVHRPGLSYG 240
Db 181 RYLGSSDKNYAKALYAKHIKIPCGGSHVETFFHEYRDNDVDFVDPHPSVHRPGLSYG 240
Qy 241 DNFAGDGNQFRYTLILCYAACAPLILELGGYIYQNCMFVNDWHASLVPVLLAAKYRP 300
Db 241 DNFAGDGNQFRYTLILCYAACAPLILELGGYIYQNCMFVNDWHASLVPVLLAAKYRP 300
Qy 301 YGYRDSRSTLVTHNLAHQGVPEASTYDGLPPEWYGALEWVPEWARRHALDKGEAVN 360
Db 301 YGYRDSRSTLVTHNLAHQGVPEASTYDGLPPEWYGALEWVPEWARRHALDKGEAVN 360
Qy 361 FLKGAVTADRIYTVSOGYSWEVTTAEGGGLNELSSRKSVLNGVINGIDINDWNPTTD 420
Db 361 FLKGAVTADRIYTVSOGYSWEVTTAEGGGLNELSSRKSVLNGVINGIDINDWNPTTD 420
Qy 421 KCLPHHYSVDDLSGKAKCAELQKELGVPREDVPLIGFIRGLDYQKIDIKMAIPELM 480
Db 421 KCLPHHYSVDDLSGKAKCAELQKELGVPREDVPLIGFIRGLDYQKIDIKMAIPELM 480
Qy 481 REDVQFVLMGSDGPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDIILLMPSRFP 540
Db 481 REDVQFVLMGSDGPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDIILLMPSRFP 540
Qy 541 CGLNQLYAMQYGTVPVYHGTGGLRDVTETFPNFGAKGEGTGWAFSPLTVDKMLWALRTA 600
Db 541 CGLNQLYAMQYGTVPVYHGTGGLRDVTETFPNFGAKGEGTGWAFSPLTVDKMLWALRTA 600
Qy 601 MSTFREHKPSWEGMLKRGMTKDHTWDHA 628
Db 601 MSTFREHKPSWEGMLKRGMTKDHTWDHA 628

RESULT 2
Q9LECO PRELIMINARY; PRT; 647 AA.
ID Q9LECO AC Q9LECO
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE STARCH SYNTHASE I-1 PRECURSOR (EC 2.4.1.21).
GN WSSI-1.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. FIELDER; TISSUE=DEVELOPING KERNELS;
RA Peng M., Hucl P., Chibbar R.N.;
RT "Isolation, characterization and expression analysis of starch
synthase I from wheat (Triticum aestivum L.).";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ292521; CAB99209.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW TRANSIT peptide; Transferase; Glycosyltransferase.
FT CHAIN 1 41
FT CHAIN 42 647
SQ SEQUENCE 647 AA; 71004 MW; E66D20010B1F9C0A CRC64;
```

Query Match 82.9%; Score 3351; DB 10; Length 647;  
Best Local Similarity 99.8%; Pred. No. 1.8e-240;  
Matches 627; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAATGVGACGLAPSVRLRADPATAARASACVVRARLRRLARGRYVAELSREGPAARPAQ 60  
|||||

```
Db 1 MAATGVGACGLAPSVRLRADPATAARASACVVRARLRRLARGRYVAELSREGPAARPAQ 60
Qy 61 QOLAPLPVPGFLAPPPAPAPQAPPTQPLPDAGYAGELAPDLLLEGIAEDSIDSIIVAAS 120
Db 61 QOLAPLPVPGFLAPPPAPAPQAPPTQPLPDAGYAGELAPDLLLEGIAEDSIDSIIVAAS 120
Qy 121 EQDSEIMDANEQPAQKVTSTVFTGEAAPYAKSGGLGVDGSLPDLAALAAHRVVMVMP 180
Db 121 EQDSEIMDANEQPAQKVTSTVFTGEAAPYAKSGGLGVDGSLPDLAALAAHRVVMVMP 180
Qy 181 RYLGSSDKNYAKALYAKHIKIPCGGSHVETFFHEYRDNDVDFVDPHPSVHRPGLSYG 240
Db 181 RYLGSSDKNYAKALYAKHIKIPCGGSHVETFFHEYRDNDVDFVDPHPSVHRPGLSYG 240
Qy 241 DNFAGDGNQFRYTLILCYAACAPLILELGGYIYQNCMFVNDWHASLVPVLLAAKYRP 300
Db 241 DNFAGDGNQFRYTLILCYAACAPLILELGGYIYQNCMFVNDWHASLVPVLLAAKYRP 300
Qy 301 YGYRDSRSTLVTHNLAHQGVPEASTYDGLPPEWYGALEWVPEWARRHALDKGEAVN 360
Db 301 YGYRDSRSTLVTHNLAHQGVPEASTYDGLPPEWYGALEWVPEWARRHALDKGEAVN 360
Qy 361 FLKGAVTADRIYTVSOGYSWEVTTAEGGGLNELSSRKSVLNGVINGIDINDWNPTTD 420
Db 361 FLKGAVTADRIYTVSOGYSWEVTTAEGGGLNELSSRKSVLNGVINGIDINDWNPTTD 420
Qy 421 KCLPHHYSVDDLSGKAKCAELQKELGVPREDVPLIGFIRGLDYQKIDIKMAIPELM 480
Db 421 KCLPHHYSVDDLSGKAKCAELQKELGVPREDVPLIGFIRGLDYQKIDIKMAIPELM 480
Qy 481 REDVQFVLMGSDGPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDIILLMPSRFP 540
Db 481 REDVQFVLMGSDGPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDIILLMPSRFP 540
Qy 541 CGLNQLYAMQYGTVPVYHGTGGLRDVTETFPNFGAKGEGTGWAFSPLTVDKMLWALRTA 600
Db 541 CGLNQLYAMQYGTVPVYHGTGGLRDVTETFPNFGAKGEGTGWAFSPLTVDKMLWALRTA 600
Qy 601 MSTFREHKPSWEGMLKRGMTKDHTWDHA 628
Db 601 MSTFREHKPSWEGMLKRGMTKDHTWDHA 628

RESULT 3
Q9SQG9 PRELIMINARY; PRT; 647 AA.
ID Q9SQG9 AC Q9SQG9
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE STARCH SYNTHASE I.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ENDOSPERM;
RA Li Z., Rahman S., Kosar-Hashemi B., Mouille G., Appels R.,
RA Morell M.K.;
RT "Cloning and characterization of a gene encoding wheat starch synthase
I.";
RL Theor. Appl. Genet. 98:1208-1216(1999).
DR EMBL; AF091803; AAD54661.1; -.
DR InterPro; IPR000531; TonB_boxC.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
SQ SEQUENCE 647 AA; 71018 MW; C5D078CCF973BD7D CRC64;
```

Query Match 82.8%; Score 3348; DB 10; Length 647;



Best Local Similarity 99.7%; Pred. No. 3e-240;  
Matches 626; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

.QY 1 MAATGVGAGCLAPSVRLRADPATAARASACVVRARLRRLARGRYVAELSGREGPAARPAQ 60  
|||||  
Db 1 MAATGVGAGCLAPSVRLRADPATAARASACVVRARLRRLARGRYVAELSGREGPAARPAQ 60

.QY 61 QOLAPPLVPGFLAPPPPPAPPAQSPAPTOPPLPDAGVGLAPDLLLEGIAEDSIDSIIVAAS 120  
|||||  
Db 61 QOLAPPLVPGFLAPPPPPAPPAQSPAPTOPPLPDAGVGLAPDLLLEGIAEDSIDSIIVAAS 120

.QY 121 EODSEIMDANEOPQAKVTRISIVFTGEAAPYAKSGGLGDCVCSLPIALAAARGHRVMVMP 180  
|||||  
Db 121 EODSEIMDANEOPQAKVTRISIVFTGEAAPYAKSGGLGDCVCSLPIALAAARGHRVMVMP 180

.QY 181 RYLNGSSDKNYAKALYTAHKKIPCFGGSHEVTFHEYRDNDVDFVDPHPSYHRPGSLYG 240  
|||||  
Db 181 RYLNGSSDKNYAKALYTAHKKIPCFGGSHEVTFHEYRDNDVDFVDPHPSYHRPGSLYG 240

.QY 241 DNFAGFQDNQFRYTLCLCYAAACEAPLLELGGYIYGQNCMFVNDWHASLVPVLLAAKYRP 300  
|||||  
Db 241 DNFAGFQDNQFRYTLCLCYAAACEAPLLELGGYIYGQNCMFVNDWHASLVPVLLAAKYRP 300

.QY 301 YGVYRDSRSTLVTHNLAHOGVEPASTYPDLGLPPEWYGALEWFPPEWARRHALDKGEAVN 360  
|||||  
Db 301 YGVYRDSRSTLVTHNLAHOGVEPASTYPDLGLPPEWYGALEWFPPEWARRHALDKGEAVN 360

.QY 361 FLKGAVVTADRIVTVSQGSWEVTTAEGGQNLNELLSSRKSVLNGIVNGIDINDWNPTTD 420  
|||||  
Db 361 FLKGAVVTADRIVTVSQGSWEVTTAEGGQNLNELLSSRKSVLNGIVNGIDINDWNPTTD 420

.QY 421 KCLPHYSVDDLSGKAKCAELQKELGLPVREDVPLIGFGRDLYQKIDLIKMAIPELM 480  
|||||  
Db 421 KCLPHYSVDDLSGKAKCAELQKELGLPVREDVPLIGFGRDLYQKIDLIKMAIPELM 480

.QY 481 REDVQFVMLGSGDPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDIILLMPSRFEP 540  
|||||  
Db 481 REDVQFVMLGSGDPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDIILLMPSRFEP 540

.QY 541 CGLNOLYAMQYCTVPVHGTGGRLDRTVETFPNPFKAGEEGTGWAFSPLTVDKMLWALRTA 600  
|||||  
Db 541 CGLNOLYAMQYCTVPVHGTGGRLDRTVETFPNPFKAGEEGTGWAFSPLTVDKMLWALRTA 600

.QY 601 MSTFREHKPSWEGLMKRGMTKDHTWDHA 628  
|||||  
Db 601 MSTFREHKPSWEGLMKRGMTKDHTWDHA 628

RESULT 4  
Q9LEB9 PRELIMINARY; PRT; 647 AA.  
AC Q9LEB9;  
DT 01-OCT-2000 (TremBLrel. 15, Created)  
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)  
DE STARCH SYNTHASE 1-2 PRECURSOR (EC 2.4.1.21).  
GN WSS1-2.  
OS Triticum aestivum (Wheat).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Triticum.  
OX NCBI\_TaxID=4565;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. FIELDER; TISSUE=DEVELOPING KERNELS;  
RA Peng M., Hucl P., Chibbar R.N.;  
RT "Isolation, characterization and expression analysis of starch  
synthase I from wheat (Triticum aestivum L.).";  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ292522; CAB99210.1; -  
DR InterPro; IPR001296; Glycos\_transf\_1.  
DR InterPro; IPR000531; TonB\_boxC.  
DR Pfam; PF00534; Glycos\_transf\_1; 1.

DR PROSITE: PS00430; TONE\_DEPENDENT\_REC\_1; UNKNOWN\_1.  
KW Transit peptide; Transferase; Glycosyltransferase.  
FT TRANSIT 1 41 POTENTIAL.  
FT CHAIN 42 647 STARCH SYNTHASE I-2.  
SQ SEQUENCE 647 AA; 71004 MW; D3374D7C3497DF6F CRC64;

Query Match 82.7%; Score 3343; DB 10; Length 647;  
Best Local Similarity 99.5%; Pred. No. 7.2e-240;  
Matches 625; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

.QY 1 MAATGVGAGCLAPSVRLRADPATAARASACVVRARLRRLARGRYVAELSGREGPAARPAQ 60  
|||||  
Db 1 MAATGVGAGCLAPSVRLRADPATAARASACVVRARLRRLARGRYVAELSGREGPAARPAQ 60

.QY 61 QOLAPPLVPGFLAPPPPPAPPAQSPAPTOPPLPDAGVGLAPDLLLEGIAEDSIDSIIVAAS 120  
|||||  
Db 61 QOLAPPLVPGFLAPPPPPAPPAQSPAPTOPPLPDAGVGLAPDLLLEGIAEDSIDSIIVAAS 120

.QY 121 EODSEIMDANEOPQAKVTRISIVFTGEAAPYAKSGGLGDCVCSLPIALAAARGHRVMVMP 180  
|||||  
Db 121 EODSEIMDANEOPQAKVTRISIVFTGEAAPYAKSGGLGDCVCSLPIALAAARGHRVMVMP 180

.QY 181 RYLNGSSDKNYAKALYTAHKKIPCFGGSHEVTFHEYRDNDVDFVDPHPSYHRPGSLYG 240  
|||||  
Db 181 RYLNGSSDKNYAKALYTAHKKIPCFGGSHEVTFHEYRDNDVDFVDPHPSYHRPGSLYG 240

.QY 241 DNFAGFQDNQFRYTLCLCYAAACEAPLLELGGYIYGQNCMFVNDWHASLVPVLLAAKYRP 300  
|||||  
Db 241 DNFAGFQDNQFRYTLCLCYAAACEAPLLELGGYIYGQNCMFVNDWHASLVPVLLAAKYRP 300

.QY 301 YGVYRDSRSTLVTHNLAHOGVEPASTYPDLGLPPEWYGALEWFPPEWARRHALDKGEAVN 360  
|||||  
Db 301 YGVYRDSRSTLVTHNLAHOGVEPASTYPDLGLPPEWYGALEWFPPEWARRHALDKGEAVN 360

.QY 361 FLKGAVVTADRIVTVSQGSWEVTTAEGGQNLNELLSSRKSVLNGIVNGIDINDWNPTTD 420  
|||||  
Db 361 FLKGAVVTADRIVTVSQGSWEVTTAEGGQNLNELLSSRKSVLNGIVNGIDINDWNPTTD 420

.QY 421 KCLPHYSVDDLSGKAKCAELQKELGLPVREDVPLIGFGRDLYQKIDLIKMAIPELM 480  
|||||  
Db 421 KCLPHYSVDDLSGKAKCAELQKELGLPVREDVPLIGFGRDLYQKIDLIKMAIPELM 480

.QY 481 REDVQFVMLGSGDPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDIILLMPSRFEP 540  
|||||  
Db 481 REDVQFVMLGSGDPIFEGWNRSTESSYKDKFRGWGFSVPVSHRITAGCDIILLMPSRFEP 540

.QY 541 CGLNOLYAMQYCTVPVHGTGGRLDRTVETFPNPFKAGEEGTGWAFSPLTVDKMLWALRTA 600  
|||||  
Db 541 CGLNOLYAMQYCTVPVHGTGGRLDRTVETFPNPFKAGEEGTGWAFSPLTVDKMLWALRTA 600

.QY 601 MSTFREHKPSWEGLMKRGMTKDHTWDHA 628  
|||||  
Db 601 MSTFREHKPSWEGLMKRGMTKDHTWDHA 628

RESULT 5  
Q9W5A3 PRELIMINARY; PRT; 643 AA.  
AC Q9W5A3;  
DT 01-OCT-2000 (TremBLrel. 15, Created)  
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)  
DE STARCH SYNTHASE I.  
GN SSI.  
OS Hordeum vulgare (Barley).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Hordeum.  
OX NCBI\_TaxID=4513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. MOREX;











```
Db 414 KCIPCHYSVDDLSGKAKKAGLQKELGLPIRPDVPLIGFIGRLDYQKGLDYLQIIPDL 473
QY 481 REDQVFMVLSGDPPIEGWMNRSTESSYKDKFRGWGFSVPVSHRITAGCDIILMPRRFP 540
Db 474 REDQVFMVLSGDPPELEDWMNRSTESIFDKFRGWGFSVPVSHRITAGCDIILMPRRFP 533
QY 541 CGLNQLYAMQYGTVPVVHGTGGLRDVTETFNPFKAGREGTGWAFSPLTVDKM 593
Db 534 CGLNQLYAMQYGTVPVVHATGGLRDVTENFPFENGEGEQGTGWAFLATENN 586

RESULT 9
Q9SBL2 PRELIMINARY; PRT; 629 AA.
ID Q9SBL2;
AC Q9SBL2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SOLUBLE STARCH SYNTHASE.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RAFIR 5765-6-1-11-3;
RA Hsieh J.S., Chen M.R., Hsing Y.I.C.;
RT "Molecular cloning of a Sorghum cDNA encoding the soluble starch
RT synthase SSSS.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF168786; AAD45815.2;
DR InterPro; IPR001296; Glycos.transf_1.
DR Pfam; PF00534; Glycos.transf_1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
SQ SEQUENCE 629 AA; 68548 MW; 24443AF594F35DE5 CRC64;

Query Match 63.4%; Score 2562.5; DB 10; Length 629;
Best Local Similarity 81.1%; Pred. No. 6.5e-182;
Matches 489; Conservative 30; Mismatches 57; Indels 27; Gaps 6;

QY 4 TGVGAGCLAPSVRLRADPATAARASACV-----VRALKR---RLARGYVAELSRREG 52
Db 5 SAVGAACL-----VLARAAAGLGLGGRGDRPRFRQVRVRRRCVAELSRREG 53

QY 53 PA--ARPAQQQLAPPLVPGFLAPPPAPAPQSPAPTOPPLPDAGVCGELAPDLLLEGTAED 110
Db 54 PAPTPRPLPALAPPLVPAFLA--PFSEPEGEPASTPPPLPDAGLGLG--LQPEGIAG 110

QY 111 SIDSIIVAASEQDSEIMDANEQPOAKVTRISIVFTGEAAPYAKSGGLGDCVGSILPTALAA 170
Db 111 SIDETVVVASEQDSEIVVGKEQARAKVTQISIVFTGEASPYAKSGGLGDCVGSILPVALLA 170

QY 171 RGRVMMVMPRYLNGSSDKNYAKALYAKIKIPCFGGSHEVTFEYRDNVDWVFVDHP 230
Db 171 RGRVMMVMPRYLNGTSDKNYAFYTEKHIRIPCFGGEHEVTFEYRDSVDWVFVDHP 230

QY 231 SYHRPGSILGDNFCAGFNDGFRYTLCLYACAEAPLILELGYLYGNCMFVNDWHASLV 290
Db 231 SYHRPGNLYGDKFAGFNDGFRYTLCLYACAEAPLVLELGYLYGNCMFVNDWHASLV 290

QY 291 PVLLAAKRYPGYVYRDSRSLVLIHNLAHQGVPEASTYPDGLCPPEWYGALEWVPEWARR 350
Db 291 PVLLAAKRYPGYVKDSRSILVLIHNLAHQGVPEASTYPDGLCPPEWYGALEWVPEWARR 350

QY 351 HALDKGAENFLKCAVYADRIYVSGYSWEVTTAGGGQGLNELLSRKSVLNGIYNGI 410
Db 351 HALDKGAENFLKCAVYADRIYVSGYSWEVTTAGGGQGLNELLSRKSVLNGIYNGI 410

QY 411 DINDWNPPTDKCLPHYSVDDLSGKAKKAGLQKELGLPIRPDVPLIGFIGRLDYQKGLD 470
```

```
Db 411 DINDWNPATDKCIPCHYSVDDLSGKAKKCSALQKELGLPIRPEVPLIGFIGRLDYQKGLD 470
QY 471 LIKMAIPELMREDVQFVMLGSGDPIPEGMMNRSTESSYKDKFRGWGFSVPVSHRITAGCD 530
Db 471 LIQIIPHLMRDQVQFVMLGSGDPELEDWMNRSTESDFDKFRGWGFSVPVSHRITAGCD 530
QY 531 ILLMPSRPEPCGLNOLYAMQYGTVPVVHGTGGLRDVTETFNPFKAGREGTGWAFSPLTV 590
Db 531 ILLMPSRPEPCGLNOLYAMQYGTVPVVHATGGLRDVTENFPFENGEGEQGTGWAFLATLT 590
QY 591 DKM 593
Db 591 ENM 593

RESULT 10
Q9FNF2 PRELIMINARY; PRT; 652 AA.
ID Q9FNF2;
AC Q9FNF2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SOLUBLE STARCH SYNTHASE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98069011; PubMed=9405937;
RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT Sequence features of the regions of 1,044,062 bp covered by thirteen
RT physically assigned pl clones.";
RL DNA Res. 4:291-300(1997).
DR EMBL; AB006701; BAB10396.1;
DR InterPro; IPR001296; Glycos.transf_1.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR002114; PTS_HPR_ser.
DR Pfam; PF00534; Glycos.transf_1.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
SQ SEQUENCE 652 AA; 72098 MW; 91E5069DCD1B2B5B CRC64;
```

```
Query Match 50.4%; Score 2038; DB 10; Length 652;
Best Local Similarity 62.9%; Pred. No. 6.2e-143;
Matches 379; Conservative 81; Mismatches 93; Indels 50; Gaps 7;
```

```
QY 43 RYVAELSRREGPAARQAQQQLAPPLVPGFLAPPP-----VLGFQLTPPGDQQTSTSTGEITHHEKKEAIDQI 111
Db 64 RFTIDAERDGS-----VLGFQLTPPGDQQTSTSTGEITHHEKKEAIDQI 111

QY 89 PLPDAGY-GELAPDLLLEGTAEDSIDSIIVAAEQDSEIMDANEQPOAKVTRISIVFTGE 147
Db 112 VMADFGVPGNRAVE---EGAAEVGIPS-----GKAEVYNNLVFTSE 150

QY 148 AAPYAKSGGLGDCVGSILPTALAAARHVRMMVMPRYLNG--SSDKNYAKALYAKIKIPCF 206
Db 151 AAPYAKSGGLGDCVGSILPTALAAARHVRMMVISPRLNGTAADKNYARAKDLGIRVTVNCF 210

QY 207 GGSHEVTFEYRDNVDWVFVDHPHSYHRPGSLYGDNFAGFNDGFRYTLCLYACAEAPLI 266
Db 211 GGSQEVSYFHEYRGDGVDFVDFHKSYPHRCNPGYDCKGAFGDNQFRFTLLCHAAEAPLV 270

QY 267 LELGGYIYQNCMFVNDWHASLVPLVLLAAKRYPGYVYRDSRSLVLIHNLAHQGVPEAST 326
Db 271 LPLGGFTYGEKSLFVNDWHAGLVPLVLLAAKRYPGYVYRDSRSLVLIHNLAHQGVPEAST 330

QY 327 YPDGLPPEWYGALEWVPEWARRHALDKGAENFLKCAVYADRIYVTVSGYSWEVTTA 386
```



```

Db 331 YTNLGLPSEWYGAVGVPTWARTHALDTGEAVNVLKGAIVTSDRIIVTSQGYAWBITTV 390
QY 387 EGGGLNELLSRKSVLNGVINGIDINWNPPTDKCLPHHYSVDLSGKAKCAELQKEL 446
Db 391 EGGYGLQDLLSRKSVINGITNGINVDWNPSTDEHIFPHYSADDSSEKIKCKMALQKEL 450
QY 447 GLPVREDVPLIFIGRLDYQKIDILKMAIPELMREDVQFVMLGSGDPFIFEGWWRSTESS 506
Db 451 GLPIRPECPMIGFICRLDYQKIDILQTAGDPLWDDIQFVMLGSGDPKYESMWRSMEEET 510
QY 507 YDKFERGWGFSVPVSHRITAGCDILLMPSRFEPCGLNQLYAMQYGVVPHVHGTGGLRDT 566
Db 511 YDKFERGWGFSVPVSHRITAGCDILLMPSRFEPCGLNQLYAMRYGTIPVPHVHGTGGLRDT 570
QY 567 VETPNPFGAKGE-ECTGWAFSPLTVDKMLWALRTAMSTPREHKPSWEGLMKRGMTKDHWTW 625
Db 571 VENFPNPAEGGAGTGTGVFTPLSKDSMVSALRLAAATYREYKQSWEGLMRGRMTNYSW 630
QY 626 DHA 628
Db 631 ENA 633

```

```

RESULT 11
Q9SEI7
ID Q9SEI7 PRELIMINARY; PRT; 575 AA.
AC Q9SEI7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE SOLUBLE STARCH SYNTHASE.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lue W.L., Wang S.M., Yu T.S., Chen J.;
RT "Characterization of Arabidopsis soluble starch synthase gene.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF121673; AAF24126.1;
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR002114; PTS_HPT_ser.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
DR PROSITE; PS00589; PTS_HPT_SER; UNKNOWN_1.
SQ SEQUENCE 575 AA; 6330 MM; 937F9F71EA842030 CRC64;

```

```

Query Match 50.1%; Score 2025.5; DB 10; Length 575;
Best Local Similarity 58.4%; Pred. No. 4.4e-142;
Matches 363; Conservative 81; Mismatches 74; Indels 13; Gaps 4;

QY 109 EDSDSIIV-----AASEODSEIMDANEQPAKVRISIVFVGEAAPYAKSGGLGD 159
Db 28 KEADQIVMADFVPGNRAVEEGAAEV--GIPSGKAEVNNLVFTVEAAPYKSGTGLGD 85
QY 160 VCGSLPTALAARGHRVNMVPRYLNG--SSDKYAKALYTAKHKIPCFGSGSHEVTFHEY 218
Db 86 VCGSLPTALAGHRVNMVSPRYLNGTAADKNYARAKDGIQVTVNCFGSGQSVSYHEH 145
QY 219 RDVNDVDFVDPHSPHRSGLYGNFGAGDNQPRYTLVYAAACEAPLILELGSYIYQNC 278
Db 146 RDGVNDVDFVDPHSPHRSGLYGNFGAGDNQPRYTLVYAAACEAPLILELGSYIYQNC 278
QY 279 MFVNDVHSLVPLVLLAAKRYGVYRDSRSLVLIHNLAHQGVPEASTYDPLGLPEWYG 338
Db 206 LFLVNDVHSLVPLVLLAAKRYGVYRDSRSLVLIHNLAHQGVPEASTYDPLGLPEWYG 338
QY 339 ALEWFPFENARRHALDKGEAVNPLKGAIVTADRIIVTSQGYAWBITTVTGGQGLNELLS 398

```

```

Db 266 AVGVWFTWARTHALDTGEAVNVLKGAIVTSDRIIVTSQGYAWBITTVTGGQGLNELLS 325
QY 399 RKSVLNGVINGIDINWNPPTDKCLPHHYSVDLSGKAKCAELQKELGLPVREDVPLIG 458
Db 326 RKSVLNGVINGIDINWNPPTDKCLPHHYSVDLSGKAKCAELQKELGLPVREDVPLIG 385
QY 459 FIGRLDYQKIDILKMAIPELMREDVQFVMLGSGDPFIFEGWWRSTESSYKDKFRGWGFS 518
Db 386 FIGRLDYQKIDILQTAGDPLWDDIQFVMLGSGDPKYESMWRSMEEETYRDKFRGWGFS 445
QY 519 VPVSHRITAGCDILLMPSRFEPCGLNQLYAMQYGVVPHVHGTGGLRDTVTFNPFQAKGE 578
Db 446 VPVSHRITAGCDILLMPSRFEPCGLNQLYAMRYGTIPVPHVHGTGGLRDTVTFNPFQAKGE 505
QY 579 -ECTGWAFSPLTVDKMLWALRTAMSTPREHKPSWEGLMKRGMTKDHWTW 628
Db 506 GAGTGWVFTPLSKDSMVSALRLAAATYREYKQSWEGLMRGRMTNYSWENA 556

```

```

RESULT 12
Q9MAC8
ID Q9MAC8 PRELIMINARY; PRT; 792 AA.
AC Q9MAC8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PUTATIVE GLYCOGEN SYNTHASE.
DE T4P13.13.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC T4P13 genomic sequence.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC008261; AAF26156.1;
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR002114; PTS_HPT_ser.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR PROSITE; PS00589; PTS_HPT_SER; UNKNOWN_1.
SQ SEQUENCE 792 AA; 87592 MM; F8C852E4119EF670 CRC64;

```

```

Query Match 31.1%; Score 1258; DB 10; Length 792;
Best Local Similarity 44.8%; Pred. No. 7.2e-85;
Matches 278; Conservative 79; Mismatches 202; Indels 62; Gaps 16;

QY 46 AELSRREGPAARPAQOQOLA-----PPLVPGFLAPP--PPAPAQSPAPTPQLPDAG--- 94
Db 181 ASVISSPVTSQPKPSDVATNGKPSVAVSSVDPPIKPESSVMTSPEKTSDFVTSFGPS 240
QY 95 ---VGLAPDLL-----LEGIAEDSDSIIVAAASEODSEIMDANEQPO--- 134
Db 241 KSRAGAFWSDPUSYLTAKAPOTSTMKTEKYVEKTPD---VASSETNEPKDEKPPPLAG 297
QY 135 AKVTRISIVFTGEAAPYAKSGGLGDVCGSLPTALAARGHRVNMVPRYLNGSSDKYAKA 194
Db 298 ANV-MNVILVAEACAPFSKGTGLGDVAGALPKSLARRGHRVNMVPRY-----AEYAEAKD 352
QY 195 LYTAHKIIPCGSGSHEVTFHEYRDNDVDFVDPHSPY-HRPGSLYGNFGAGDNQPRY 253
Db 353 LGVRRYRYKA--GQDMVEMYHAFIDGVDFVDFIDSPFDRHLSNNIYGCN---RLDILARM 407
QY 254 TLLCYAAACEAPLILELGSYIYQNC-MCEVNDVHSLVPLVLLAAKRYGVYRDSRSLV 312
Db 408 VLFCKAAVEVPWYPCGGVCGYGDGNLAFIANDWHTALLPVYLKAYIRDRHINKYTRSLV 467

```



[illegible]

```

RESULT 13
O65365 PRELIMINARY; PRT; 630 AA.
ID O65365
AC O65365;
DT 01-AUG-1998 (T-EMBLrel. 07, Created)
DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE STARCH SYNTHASE.
OS SSSI.
GN Ipomoea batatas (Sweet potato) (Batate).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Convolvulaceae; Ipomoea.
RX NCBI_TaxID=4120;
RN [1]
RP SEQUENCE FROM N.A.
RA Harn C.H., Bae J.M., Liu J.R.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF068834; AAC19119.1; -
DR Mendel; 29714; Ipoba:Sssl;29714.
DR InterPro; IPR001296; Glycos.transf_1.
DR InterPro; IPR002114; PTS_HPr_Ser.
DR Pfam; PF00534; Glycos.transf_1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
DR SEQUENCE 630 AA: 69961 MW: C55669D761B1A12A CRC64;

```

	Query Match	30.88;	Score 1245;	DB 10;	Length 630;
	Best Local Similarity	44.88;	Pred. No. 4.7e-84;		
	Matches 275;	Conservative 82;	Mismatches 197;	Indels 60;	Gaps 14;
Qy	52	GPAARPAQOOQ--LAPPLVPGFLA---PPPPAPAQSPAPTQ-PPPL---PDAGVGGELAPDL	102		
Db	25	GSTERPAVKSKVLSLPSQVTSSTVNSQEPAAKATVETVKSSQVMPDLVDVSQKNVTFSFSREI	84		
Qy	103	LLEGTAEDSIDSIIVAAEQSDSEIMDANEQQAQKV-----TRSI	141		
Db	85	LL---SEYFPSSSMASASTLEDEKRDHRESSKEIDVCTEDVPNBDLKPPLAGTNVMNV	141		
Qy	142	VFVTEGAAPYAKSGGLGDVCGSLPTALAARGHRVNMVPRYLNGSSDKNYAKALYIAKHI	201		
Db	142	ILVCAEAPAKWSTGGGLGDVAGALPKALARGRHVRNMVVPVLY-----GNVAPQHTGVYRK	195		
Qy	202	KIPCEGGGSHVETTFHEYRDNDVDMVEVDHPSY--HRPGSLYGDNGFAGNQRFYTLICVAA	260		
Db	196	MFKIDQQDMENVNFIAYINDVDFIDTSPDFIGHRGNNIYVGGN---RVDLIKRMDLFCRAA	252		

Qy	261	CEAPLILELGGYIGQ -NCMFVVNDWHASLVPVLLAAKYPYGVYVRDSDRSTLVHNLHAHQ	312
Db	253	IYVPNHPVPCGGTCYGDGNLVFTANDWHMTALLPVYLKAYFRDNGVWKFRSLVLIHINIAHQ	312
Qy	320	GVPEASTYDPLGLUPPBYWGALEWVPEWARRHALDKGEAVNPLKGAVVYADTRIVTSQGY	379
Db	313	GRCPMDDFSIVDLPAQYADLFLKYDPV-----GGDHFNIFAAGLKTADRVVTVSHGY	364
Qy	380	SWEVTTAAGGQGNELLSRKSVLNGVINGIDINDWNPTTKCLPH-----HYSVDLL-SG	434
Db	365	AWEKTSSEGGWGLNGLRNENKLOGVINGIDIEEWNPOLDYVLYLKSDBGYANSYSLDTLQTG	424
Qy	435	KAKCKAELOKELGVPREDVPLIGFTGRIDYOKGIDLIKMAIPELMREDVQVFMVLSGDDP	494
Db	425	KPOCKAALQKEMMLPVRDDVPLIGFTGRIDYOKGVDLIAEATPMMGQDVLVMLGTGRP	484
Qy	495	IFEGWNRSTESSYKDKFRGVGVSVPVSHRITAGCDILLMPSEPCGLNQLYAMQYGTIV	554
Db	485	DLEQMLQIBEQYGDYKVRGVWGVSVKTAHRTAGADILLMPSEPCGLNQLYAMSYGTV	544
Qy	555	PVHVTGGLRLDRTVETNPFCAKGEECTGNWFSPLTVDKMLMALRTAMSTFRHKPQSWEG	614
Db	545	PVHAVGGLRLDTVQPPDPF---NESGYGWTFFGRAEANQLIDALGNCLLTYRQYKQSWEG	601
Qy	615	MKRGMTKDHWTWHA	628
Db	602	QRRGMQDLSWDHA	615
RESULT 14			
Q48899	PRELIMINARY; PRT; 732 AA.		
ID	O48899		
AD	O48899		
DT	01-JUN-1998 (T-EMBLrel. 06, Created)		
DT	01-JUN-1998 (T-EMBLrel. 06, Last sequence update)		
DT	01-JUN-2001 (T-EMBLrel. 17, Last annotation update)		
DE	STARCH SYNTHASE ISOFORM 2STII-1 (BC 2.4.1.21) (STARCH (BACTERIAL		
DE	COGEN) SYNTHASE) (GLYCOGEN SYNTHASE) (ADP-GLUCOSE-6-PHOSPHATE		
DE	GLUCOSYLTRANSFERASE) (FRAGMENT).		
GN	SSSI OR ZSSIIA.		
OS	Zea mays (Maize).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade		
OC	Panicoidae; Andropogoneae; Zea.		
OC	NCBI_TaxID=4577;		
RL	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=W61A; TISSUE=ENDOSPERM;		
RA	Knight M.E., Harrn C., Lilley C.E.R., Guan H., Singletary G.W.,		
RA	Mu-Forester C., Wasserman B.P., Keeling P.L.;		
RL	Plant J. 0:0-0(1998).		
RL	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=W61A; TISSUE=ENDOSPERM;		
RA	Harrn C., Knight M., Ramakrishnan A., Guan H., Keeling P.L.,		
RA	Wasserman B.P.;		
RL	Plant Mol. Biol. 0:0-0(1998).		
CC	- CATALYTIC ACTIVITY: ADP-GLUCOSE + (1,4-ALPHA-D-GLUCOSYL)(N) - ADP		
CC	+ (1,4-ALPHA-D-GLUCOSYL)(N-1).		
DR	EMBL; AF019296; AAD13341.1; -		
DR	Mendel; 24658; Zeama; SSSI; 24658.		
DR	InterPro; IPR001296; Glycos_transf_1.		
DR	InterPro; IPR002086; Aldehyde_dehydr.		
DR	Pfam; PF00534; Glycos_transf_1.		
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.		
KW	Transferase; Glycosyltransferase.		
FT	NON_TER 732		
SQ	SEQUENCE 732 AA; 29A6745055090F3B CRC64;		

Query Match 30.4%; Score 1227.5; DB 10; Length 732;  
Best Local Similarity 40.3%; Pred. NO. 1.2e-82;  
Matches 276: Conservative 104; Mismatches 212; Indels 93;











OM of: US-09-674-824-2 to: EST:\* out\_format : pfs  
 Date: Mar 28, 2002 4:55 PM  
 About: Results were produced by the GenCore software, version 4.5,  
 Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
 -O=/cgn2\_1/USPTO.spool/US09674824/runat\_27032002\_094250\_18407/app\_query.fasta\_1.824  
 -DB=EST -FAST=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000  
 -MINMATCH=0.100 -LOPC=1.000 -LOPEXT=0.000 -QGAPOP=4.500  
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40 cdi  
 -LIST=45 -DCCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0  
 -ALIGN=15 -MODE=LOCAL -OUTFM=pfs -NORM=ext -HEAPEXT=500  
 -MINLEN=0 -MAXLEN=200000000 -USER=US09674824\_CGNI\_1\_3637  
 -NCPU=6 -ICPU=3 -LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30  
 -NO\_XLPXY -WAIT -THREADS=1

## Search information block:

Query: US-09-674-824-2  
 Query length: 756  
 Database: EST:\*  
 Database sequences: 11351937  
 Database length: 1077921985  
 Search time (sec): 1499.200000

## score.list:

Sequence	Strd Orig	zScore	EScore	Len	Documentation
gb_est2:BG343782	+ 1120.00	1305.81	1.4e-63	988	! BG343782 HVSMeg0006L20f Hordeum
gb_est1:AV832913	- 944.00	1103.12	2.8e-52	706	! AV832913 AV832913 K. Sato unpub
gb_est1:BA418354	+ 924.00	1076.68	8.3e-51	988	! BA418354 SCU031.D08R990531 ITEC
gb_est2:BF113316	+ 842.00	985.41	1.0e-45	596	! BF113316 EST440906 tomato break
gb_est1:BP434991	+ 807.00	944.62	1.9e-43	588	! BP434991 EST406069 tomato break
gb_est1:AW678087	+ 789.00	923.99	2.7e-42	562	! AW678087 WSL_13 G10.bl.A002 Wat
gb_est1:AW932199	+ 735.50	861.41	8.1e-39	564	! AW932199 EST358042 tomato fruit
gb_est1:AV832689	+ 710.00	833.57	2.9e-37	454	! AV832689 AV832689 K. Sato unpub
gb_est2:BF482660	+ 682.00	801.81	1.7e-35	408	! BF482660 WHE2301-2304_E13_E1329
gb_est1:AW759569	+ 643.50	753.03	8.8e-33	618	! AW759569 si45a09.y1 Gm-cl027 G1
gb_est1:AW59946	+ 640.50	749.45	1.4e-32	623	! AW59946 si56b01.y1 Gm-cl027 G1
gb_est1:AW599325	+ 638.50	747.57	1.8e-32	592	! AW599325 gb13e02.y1 Moss EST
gb_est1:AU075475	+ 613.00	721.41	5.1e-31	336	! AU075475 AU075475 Rice cDNA fr
gb_est1:AW438020	+ 555.00	653.27	3.2e-27	411	! AW438020 ST83F01 Pine Triplex
gb_est2:BG607242	+ 543.00	639.47	1.9e-26	401	! BG607242 WHE2493.B06_D1125 Trif
gb_est1:BE345184	+ 535.50	627.71	8.4e-26	557	! BE345184 NF09I081F1073 Devel
gb_est1:BE347900	+ 517.50	605.98	1.4e-24	601	! BE347900 SP08603.y1 Gm-cl041 G1
gb_est2:BG350240	+ 502.50	588.09	1.4e-23	625	! BG350240 AV441053 Arabidopsis t
gb_est1:AV524142	+ 496.00	578.03	4.3e-23	734	! AV524142 AV524142 Arabidopsis t
gb_est2:BG599615	+ 495.50	580.07	3.8e-23	614	! BG599615 EST504510 cSTS Solanum
gb_est2:BG406951	+ 494.00	576.75	5.8e-23	729	! BG406951 186E09 Mature tuber la
gb_est2:BG3511920	+ 485.00	565.45	2.5e-22	794	! BG351192 134D01 Mature tuber la
gb_est1:BG351175	+ 484.00	565.25	2.5e-22	696	! BG351175 107B03 Mature tuber la
gb_est1:AV441053	+ 473.50	552.25	1.3e-21	773	! AV441053 AV441053 Arabidopsis t
gb_est1:AW569438	+ 468.50	548.29	2.3e-21	635	! AW569438 si87a06.y1 Gm-cl031 G1
gb_est2:BF424284	+ 463.50	543.69	4.0e-21	548	! BF424284 sr36f10.y1 Gm-cl051 G1
gb_est1:BE414362	+ 460.00	537.44	9.0e-21	695	! BE414362 SCU008.H03 R990714 Tr
gb_est2:BG456652	+ 459.00	535.53	1.0e-20	681	! BG456652 NF083D09PL1F1076 Phos
gb_est2:BG449742	+ 455.50	532.76	1.6e-20	754	! BG449742 088A10 Mature tuber la
gb_est2:BG350306	+ 448.50	530.63	2.2e-20	652	! BG350306 089D01 Mature tuber la
gb_est2:BI131080	+ 446.00	520.29	8.1e-20	335	! BI131080 C19776 Rice panicle at r
gb_est1:AV418330	+ 443.50	517.28	1.2e-19	765	! AV418330 EST5311930 GESD Medica
gb_est2:BI406367	+ 439.00	520.71	7.7e-19	294	! BI406367 167H04 Lotus japonic
gb_est2:BG600852	+ 436.50	507.80	4.0e-19	882	! BG600852 EST505747 cSTS Solanum
gb_est2:BI406798	+ 434.50	507.45	4.2e-19	709	! BI406798 180H08 Mature tuber la
gb_est2:BI406287	+ 433.00	506.50	4.8e-19	649	! BI406287 159B05 Mature tuber la
gb_est2:BI406287	+ 432.50	505.50	5.4e-19	679	! BI406287 159B05 Mature tuber la

gb\_est1:AW472190 + 426.00 499.73 1.1e-18 555 ! AW472190 si21b02.y1 Gm-cl029  
 gb\_est2:BF460244 + 419.00 488.78 4.6e-18 753 ! BF460244 073C04 Mature tuber  
 gb\_est1:AV638089 + 416.50 491.08 3.4e-18 424 ! AV638089 AV638089 Chlamydomo  
 gb\_est2:BG652154 + 415.50 488.02 5.1e-18 522 ! BG652154 sad74h10.y1 Gm-cl05  
 gb\_est2:BG350735 + 414.00 484.12 8.4e-18 661 ! BG350735 097D10 Mature tuber  
 seq\_name: gb\_est2:BG343782  
 seq\_documentation\_block:  
 LOCUS BV343782 988 bp mRNA EST 29-MAY-2001  
 DEFINITION HVSMeg0006L20f Hordeum vulgare pre-anthesis spike EST library  
 HVCNAA0008 (white to yellow anther) Hordeum vulgare cDNA clone  
 HVSMeg0006L20f, mRNA sequence.  
 ACCESSION BV343782  
 VERSION BV343782.1 GI:13156111  
 KEYWORDS EST.  
 SOURCE barley.  
 ORGANISM Hordeum vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
 ; Triticeae; Hordeum.  
 REFERENCE 1 (bases 1 to 988)  
 AUTHORS Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu  
 ,Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo  
 ,T., Saski,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and  
 Wood,T.  
 TITLE Development of a genetically and physically anchored EST resource  
 for barley genomics  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Wing RA  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: rwing@clemson.edu  
 Seq primer: AATTACCCCTCACTAAGGG  
 High quality sequence stop: 845.  
 FEATURES  
 source  
 1. 988  
 /organism="Hordeum vulgare"  
 /cultivar="Morex"  
 /db\_xref="taxon:4513"  
 /clone="HVSMeg0006L20f"  
 /clone\_lib="Hordeum vulgare pre-anthesis spike EST library  
 HVCNAA0008 (white to yellow anther)"  
 /tissue\_type="pre-anthesis spike"  
 /lab\_host="SOLR"  
 /note="Vector: lambdaZAP; Site\_1: EcoRI; Site\_2: XhoI; For  
 more details on library preparation and sequence analysis  
 see http://www.genome.clemson.edu/projects/barley/ To  
 order a clone see http://www.genome.clemson.edu/orders"  
 BASE COUNT 243 a 223 c 291 g 211 t 20 others  
 ORIGIN  
 alignment\_scores:  
 Quality: 1120.00 Length: 256  
 Ratio: 4.786 Gaps: 5  
 Percent Similarity: 91.406 Percent Identity: 88.281  
 alignment\_block:  
 US-09-674-824-2 x BG343782 ..  
 Align seg 1/1 to: BG343782 from: 1 to: 988  
 504 GluSerSerTyrLysAspLysPheArgGlyTrpValGlyPheSerValPr 520  
 |||||  
 3 GAGTCGAGCTACAAGGATAAATTCCTCGGATGGTGGATTAGTTGTTCC 52  
 520 oValSerHisArgIleThrAlaGlyCysAspIleLeuLeuMetProSera 537  
 |||||  
 53 AGTTTCCACAGATAACTCGAGGTGCGATATATTGTTAATGCCATCCA 102



537 rgPheGluProCysGlyLeuAsnGlnLeuThrAlaMetGlnTyrGlyThr 553  
|||||  
103 GATTGAACCTTCGGCTCTTAATCAGCATATATGCAATATGGTACA 152  
554 ValProValValHisGlyThrGlyGlyLeuArgAspThrValGluThrPh 570  
|||||  
153 GTTCCTCTGTTCATGAACTGGGGCTTCGAGACACGGTGAGACCTT 202  
570 eAsnProPheGlyAlaLysGlyGluGluGlyThrGlyThrAlaPheSerP 587  
|||||  
203 CAACCCCTTTGGTCAAAAGAGAGAGGCTACAGGCTGGCGCTTCTCAC 252  
587 roLeuThrValAspLysMetLeuThrAlaLeuArgThrAlaMetSerThr 603  
|||||  
253 CACTAACCGTGGAAAAATGTTGGGCATTCGGAACCGCATTTCCAGC 302  
604 PheArgGluHisLysProSerTrpGluGlyLeuMetLysArgGlyMetTh 620  
|||||  
303 TTTAGGAGCACACAGCCCTCTGGAGGGCTCATGAAGCGAGGCATGAC 352  
620 rLysAspHisThrTrpAspHisAla . ProSerSerThrSerArgSerSer 636  
|||||  
353 GAAAGACCATACGTTGGGACCATGCGCGCGAGCAGTACGAGCAGATCTCG 402  
637 SerGlyProSerTrpThrAsnProThrSerCysArgArgGlyLeuGlyAr 653  
|||||  
403 AGTGGGCTTCGTGGACCAACCTACGTATGATAGCTGGGACTGGGAG 452  
653 gSerLysCysGluSerProSerAlaLeuLysThrSerSerSerPheA 670  
|||||  
453 GTCGAAGCGCCAGCTCTCTGAGCTCTCAAGACATGTTCTCTCATCTCC 502  
670 rgGlyProGlu... GlyTyrProCysThrLeuArgCysProAlaThrVal 685  
|||||  
503 CGGGCCGGAAGAGGATACCCCTGTACATTCGGTTGCTGCTACAGTA 552  
686 GluSerGlnCysAlaCysLeuLeuTrpPheAlaGlySerArgThrTyrAs 702  
|||||  
553 GAGTCGCATAGCCCTTGCTGCT . TGTTCCCGGTTCCGAGATGACGA 601  
702 pGlyCysAlaAlaAlaValThrAlaSerGlyGlyArgGln . LeuGln 718  
|||||  
602 CGGTGCAACGCTGCGCG . . GCAGCTTCGGGTGGATCACANNNTACAG 648  
719 PheTrpGlyIleArgLysGlyCysAlaAlaGlyTrpLeuThrAlaLysHi 735  
|||||  
649 TTTGGGGAATAGGAAGGGAGTCTCGGTGGGTGTACAGCTGAATCAGAAAC 698  
735 shiSerAspGlySer . LeuSerValArgValThrAlaGluIleArgAsn 751  
|||||  
699 CACCTCAGTCGACGCTCTCTCGGTGGGTGTACAGCTGAATCAGAAAC 748  
752 GlnLeuValThr 755  
|||||  
749 CCACCTGGTGTCT 760

seq\_name: gb\_est1:AV832913

seq\_documentation\_block: 706 bp mRNA EST 22-JUN-2001  
LOCUS AV832913 K. Sato unpublished cDNA library: Hordeum vulgare subsp.  
DEFINITION AV832913 K. Sato unpublished cDNA library: Hordeum vulgare subsp.  
vulgare top three leaves adult, heading stage Hordeum vulgare  
subsp. vulgare cDNA clone rbaal2b23, mRNA sequence.

ACCESSION AV832913  
VERSION AV832913.1 GI:14525002  
KEYWORDS EST.  
SOURCE Hordeum vulgare subsp. vulgare.  
ORGANISM Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
1 Triticeae; Hordeum.  
REFERENCE 1 (bases 1 to 706)

AUTHORS  
TITLE  
JOURNAL  
COMMENT

Sato, K.  
Barley EST sequencing project in NIG and Okayama Univ  
Unpublished (2001)  
Contact: Kazuhiro Sato  
Research Institute for Bioresources  
Okayama University, Barley Germplasm Center  
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan  
Email: kassato@rib.okayama-u.ac.jp,  
URL: http://www.rib.okayama-u.ac.jp/barley/  
Sato, K., Saisho, D., Takeda, K., Shini, T. and Kohara, Y. Direct  
submission:  
database: http://www.shigen.nig.ac.jp/barley/Barley.html.  
Location/Qualifiers

FEATURES  
source

1. 706  
/organism="Hordeum vulgare subsp. vulgare"  
/cultivar="Haruna Nijo"  
/db\_xref="taxon:112509"  
/clone="rbaal2b23"  
/clone\_lib="K. Sato unpublished cDNA library: Hordeum  
vulgare subsp. vulgare top three leaves adult, heading  
stage"  
/tissue\_type="top three leaves"  
/dev\_stage="adult, heading stage"  
174 a 213 c 162 g 155 t 2 others

BASE COUNT  
ORIGIN

alignment\_scores:  
Quality: 944.00 Length: 214  
Ratio: 4.744 Gaps: 4  
Percent Similarity: 92.991 Percent Identity: 90.187

alignment\_block:  
US-09-674-824-2 x AV832913/rev ..  
Align seg 1/1 to reverse of: AV832913 from: 1 to: 706

546 LeuTyrAlaMetGlnTyrGlyThrValProValHisGlyThrGlyGlu 562  
|||||  
705 CTATATGCTATGCAATATGGTACAGTTCCTCTGTTTCATCGAACTGGGG 656  
562 YLeuArgAspThrValGluThrPheAsnProPheGlyAlaLysGlyGluG 579  
|||||  
655 CCTTCGAGACACGGTGAGACCTTCAACCTTTTGGTGCAGAAAGGAGAGG 606  
579 luGlyThrGlyTrpAlaPheSerProLeuThrValAspLysMetLeuTrp 595  
|||||  
605 AGGCTACAGGGTGGCGTTCTCACACTAACCGTGGAAAAAATGTTGTGG 556  
596 AlaLeuArgThrAlaMetSerThrPheArgGluHisLysProSerTrpGlu 612  
|||||  
555 GCATTGGCAACCGCATTTTCACCGTTTAGGAGGACACAGCCCTCTCTGGGA 506  
612 uGlyLeuMetLysArgGlyMetThrLysAspHisThrTrpAspHisAla . 628  
|||||  
505 GGGGCTCATGAAGCGAGCATGACGAAAGACCATACGTGGGACCATGCGC 456  
629 ProSerSerThrSerArgSerSerSerGlyProSerTrpThrAsnProTh 645  
|||||  
455 CCGAGCATACGACGACATCTTCGAGTGGGCTTCGTTGGGACCAACCTTAC 406  
645 rSerCysArgArgGlyLeuGlyArgSerLysCysGluSerProSerAlaL 662  
|||||  
405 GTCATGTAGACTGGGACTGGGAGGTCCGAGGCCAGCTCTCTTTCAGCTC 356  
662 euLysThrSerSerSerPheArgGlyProGlu... GlyTyrProCys 677  
|||||  
355 TGAAGACATGTTCTCTCATCTTCGCGCGCGGAGAGGATACCCCTGT 306  
678 ThrLeuArgCysProAlaThrValGluSerGlnCysAlaCysLeuLeuTr 694  
|||||  
305 ACATTGCGGTGTCTGCTACAGTAGAGTCGCANTCGCGTTCGTTGCT.TG 257



```

694 pPhealaGlySerArgThrTyrrAspGlyCysAlaAlaAlaValThrA 711
|||||
256 GTTCGCCGTTTCGAGAGTAGACGAGCGTTCGCAACGCTGCCGG...GCAG 210
|||||
711 laSerGlyGlyArgGlnLeuGlnPheThrPglIleArgLysGlyCysAla 727
|||||
209 CTTCCGGTGGGATGACAGTAGTACAGTTTGGGGAAATAAGGAGGAGTGCT 160
|||||
728 AlaGlyTrpLeuThrAlaLysHisHisSerAspGlySer..LeuSerVal 743
|||||
159 GCAGAGTGTTCGACAGCANAACACACCTACAGTCGCAGCCTCTCTCGGTG 110
|||||
744 ArgValThrAlaGluIleArgAsnGlnLeuValThrLeu 756
|||||
109 GGTGTTACAGCTGAATCAGAAACCACTGGTGCTCTT 71
|||||
seq_name: gb_est1:BE418354

seq_documentation_block: 988 bp mRNA EST 24-JUL-2000
LOCUS BE418354
DEFINITION SCL031.D08R90531 ITEC SCL Wheat Leaf Library Triticum aestivum
cDNA clone SCL031.D08, mRNA sequence.
ACCESSION BE418354
VERSION BE418354.1 GI:9416200
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 988)
REFERENCE Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
Sorrenti,M., Warburton,M. and Wenzel,G.
International Triticale EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticeae
Unpublished (2000)
Contact: Cloutier S
Cereal Research Centre, Agriculture & Agri-Foods Canada
Winnipeg MT CANADA
Tel: 204 983 2340
Fax: 204 983 4604
Email: scloutier@em.agr.ca
International Triticale EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.
Location/Qualifiers
1. 988
/organism="Triticum aestivum"
/cultivar="Thatcher Lrl"
/db_xref="taxon:4565"
/clone="SCL031.D08"
/clone_lib="ITEC SCL Wheat Leaf Library"
/tissue_type="young leaf"
/dev_stage="after 24 hour challenge with LR-AVR1"
/notes="vector: Lambda ZAP; 1.0 kbp average insert size."
BASE COUNT 227 a 199 c 244 g 247 t 71 others
ORIGIN

alignment_scores:
Quality: 924.00 Length: 225
Ratio: 4.738 Gaps: 9
Percent Similarity: 86.667 Percent Identity: 84.889

alignment_block:
US-09-674-824-2 x BE418354 ..
Align seq 1/1 to: BE418354 from: 1 to: 988
475 AlaIleProGluLeuMetArgGluAspValGlnPheValMetLeuGlyse 491
|||||
10 GCCATTCAGAGCTCATGAGGAGGAGCGTCAATTTGTCTCATGCTTGATC 59
|||||
491 rGlyAspProIlePheGluGlyTrpMetArgSerThrGluSerSerThrL 508
|||||
60 TGGGATCCCAATTTTGAAGGCTGGATGATATCTACCGAGTCGAGTTTACA 109
|||||
508 ysAspLysPheArgGlyTrpValGlyPheSerValProValSerHisArg 524
|||||
110 AGGATAAATTCCTGGATGGTGGATTTAGTGTTCAGTTTCCACAGATTGAA 159
|||||
525 IleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGluProCy 541
|||||
160 ATAACGTCAGGTGGGATATATTGTTAATGCCATCCAGAGATTGAACTTG 209
|||||
541 sGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThr.ValProValVal 557
|||||
210 CGCTCTTAATCAGCTATATGCTATGCAATATGTCACAGTTCTCTGTAGT 259
|||||
558 HisGlyThrGlyGlyLeuArgAspThrValGluThrPheAsnProPheGl 574
|||||
260 CATGGAACCTGGGGGCTCCGAGACACAGTCGAGACCTTCAACCTTTTGG 309
|||||
574 yAlaLysGlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrValA 591
|||||
310 TGCAAAAGGAGAGAGGGGTACAGGGTGGGGTTCACCGCTAACCGTGG 359
|||||
591 sPlysMetLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHis 607
|||||
360 ACAAGATGTTGTGGGCATTTCGAACCGCATGTGCGACATTCAGGGAGCAC 409
|||||
608 LysProSerTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisTh 624
|||||
410 AGCCCTTCTGGAGGGGCTNATNAAGCGAGGCGATGACGAAAGACCATAC 459
|||||
624 r.TrpAspHisAla..ProSerSerThrSerArgSerSerGlyProSe 640
|||||
460 CGTGGGACCATCGCCCGAGCAGTAGCAGACAGATNTTTCAGTGGGCTTC 509
|||||
640 rTrpThrAsnProThrSer.CysArgArgGlyLeu.GlyArgSerLysCy 656
|||||
510 GTGGACCAACCTTACNTCATTTAGACNGGGACTGGGGAGNGTCCAAGTG 559
|||||
656 sGluSerProSer.AlaLeuLysThrSer.SerSerSerPhe...ArgGl 671
|||||
560 CGAGTCTCTTTGAGCTTTGAANACATCCCTCTTATTCTTTTCGCGGNC 609
|||||
671 yProGluGlyTyrrProCysThrLeuArgCysProAlaThrValGluSerG 688
|||||
610 CCGGGAAGGATACCCCTGTNCATT..... 634
|||||
688 lncysalacys 691
|||||
635 ..TGCCTTGT 643

seq_name: gb_est2:BF113316
seq_documentation_block: 596 bp mRNA EST 18-MAY-2001
LOCUS BF113316
DEFINITION EST440906 tomato breaker fruit Lycopersicon esculentum cDNA clone
cLEGA4E24 5' sequence, mRNA sequence.
ACCESSION BF113316
VERSION BF113316
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 596)
REFERENCE Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,

```







[illegible][illegible]

```

is 20.
Seq primer: JEN REV
High quality sequence start: 17
High quality sequence stop: 553
POLYA=NO.

FEATURES             Location/Qualifiers
     source            1..562
                        /organism="Sorghum bicolor"
                        /db_xref="taxon:4558"
                        /clone_lib="Water-stressed 1 (WS1)"
                        /note="Organ: Mix of 5-week old plants on days 7 & 8 after
                        water was withheld; Vector: Lambda Zap; Site 1: XhoI;
                        Site 2: EcoRI; The library was made from poly-A RNA in the
                        cloning vector lambda Zap II. Clones to be sequenced were
                        prepared by mass excision."
     BASE COUNT       153 a      126 c      142 g      140 t      1 others
     ORIGIN

alignment_scores:
    Quality: 789.00      Length: 185
    Ratio: 4.561        Gaps: 1
    Percent Similarity: 93.514      Percent Identity: 77.297

alignment_block:
US-09-674-824-2 x AW678087      ..

Align seg 1/1 to: AW678087 from: 1 to: 562

461 GlyArgLeuAspTyrGlnLysGlyIleAspLeuIleLysMetAlaIlepr 477
|||||:.....:|||||:|||||:|||||:|||||:|||||
9 GGNAAGATTGGATTATCAGAAGCATTCATCTCAATTCATCATACC 58

477 cgluLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAspp 494
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||
59 ACATCTCATCGGGACGACGTTCAATTGTGTCATGCTTGGATCTGGTGACC 108

```







/tissue\_type="top three leaves"  
/dev\_stage="adult, heading stage"  
BASE COUNT 112 a 84 c 126 g 130 t 2 others  
ORIGIN

alignment\_scores:  
Quality: 710.00 Length: 143  
Ratio: 5.035 Gaps: 0  
Percent Similarity: 98.601 Percent Identity: 95.105

alignment\_block:  
US-09-674-824-2 x AV832689 ..

Align seg 1/1 to: AV832689 from: 1 to: 454

106 GlyIleAlaGluAspSerIleAspSerIleValAlaAlaSerGluGI 122  
|||||  
27 GGGATGCTGAGGATCTCTC.GACACCATAGTTGTGGCTGCAAGTGAGCA 75  
122 nAspSerGluIleMetAspAlaAsnGluInProGlnAlaLysValThrA 139  
76 GGATCTCAGATCATGGATGCCACGACCAACCTCTAGCTAAAGTTACAC 125  
139 rgSerIleValPheValThrGlyGluAlaAlaProTyrAlaLysSerGly 155  
|||||  
126 GTAGCATCGTGTGTGACTGGTGAAGCTGCTCTTATGCAAGTCAGGG 175  
156 GlyLeuGlyAspValCysGlySerLeuProIleAlaLeuAlaAaGGL 172  
|||||  
176 GGGCTGGGAGATGTTGTGTTCTGCTGCCAATGCTCTTGTCTCGTGG 225  
172 yHisArgValMetValValMetProArgTyrTyrLeuAsnGlySerSerAspL 189  
|||||  
226 TCACCGTGTGATGTTGTTAATGCCGAGATACTTAAATGGGACCTCTCAT 275  
189 yAsnTyrAlaLysAlaLeuTyrThrAlaLysHisIleLysIleProCys 205  
|||||  
276 AAAACTATGCAAGGCAATTATACCCGGGAAGCACATTAAGATTCCATGC 325  
206 PheGlyGlySerHisGluValThrPheHisGluTyrArgAspAsnVa 222  
|||||  
326 TTCGGGGGATCACATGAAGTCACCTTTTTCATGAGTATAGACACACGT 375  
222 LAspTrpValPheValAspHisProSerTyrHisArgProGlySerLeut 239  
|||||  
376 CGATTTGGGTGTTTGTGATCATCATCATATATATAGACACCAAGTTGT 425  
239 yrGlyAspAsnPheGlyAlaPheGlyAsp 248  
|||||  
426 ATGGAGATAATTTTGGTCTTTTGGTGAT 454

seq\_name: gb\_est2:BF482660

seq\_documentation\_block:  
LOCUS BF482660 408 bp mRNA EST 06-DEC-2000  
DEFINITION WHE2301-2304\_E13\_ZS Wheat pre-anthesis spike cDNA library  
Triticum aestivum cDNA clone WHE2301-2304\_E13\_E13, mRNA sequence.  
ACCESSION BF482660  
VERSION BF482660.1 GI:11565884  
KEYWORDS EST.  
SOURCE bread wheat.  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae  
; Triticeae; Triticum.  
1 (bases 1 to 408)  
Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han  
, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,  
Seaton, C.L. and Tong, J.C.  
The structure and function of the expressed portion of the wheat  
genomes - Pre-anthesis spike cDNA library  
Unpublished (2000)

## COMMENT

Contact: Olin Anderson  
US Department of Agriculture, Agriculture Research Service, Pacific  
West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA  
Tel: 5105595773  
Fax: 5105595818  
Email: oandersn@pw.usda.gov  
Sequence have been trimmed to remove vector sequence and low  
quality sequence with phred score less than 20  
Seq primer: Stragene SK primer.

## FEATURES

## Source

1..408  
/organism="Triticum aestivum"  
/cultivar="Chinese Spring"  
/db\_xref="taxon:4565"  
/clone="WHE2301-2304\_E13\_E13"  
/clone\_lib="Wheat pre-anthesis spike cDNA library"  
/tissue\_type="Spike before anthesis"  
/dev\_stage="Adult plant"  
/lab\_host="E. coli SOLR"  
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
Site\_1: EcoRI; Site\_2: XhoI; Plants were grown in the  
greenhouse. Whole spike with awns trimmed, white, green  
and yellow anther were collected and total RNA, and  
poly(A) RNA were prepared, a cDNA library was made, and  
the cDNA clones were in vivo excised to give phagescript  
phagemids in the TJ Close lab (Choi, Close, Fenton) at  
the University of California, Riverside. Plasmid DNA  
preparations and DNA sequencing were performed in the OD  
Anderson lab (all other authors)."

BASE COUNT 108 a 76 c 99 g 125 t  
ORIGIN

## alignment\_scores:

Quality: 682.00 Length: 136  
Ratio: 5.128 Gaps: 0  
Percent Similarity: 97.794 Percent Identity: 94.118

## alignment\_block:

US-09-674-824-2 x BF482660 ..

Align seg 1/1 to: BF482660 from: 1 to: 408

427 TyrSerValAspAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGI 443  
1 TATTTGTCGATGACCTCTCTGGAAGGCCAAATGTAAGCTGAATTGCA 50  
443 nLysGluLeuGlyLeuProValArgGluAspValProLeuIleGlyPheI 460  
:|||||  
51 CAAGGAGCTGGGTTTACCTGTAAAGGAGGATGTTCTCTGATTTGCTTTA 100  
460 leGlyArgLeuAspTyrGlnLysGlyIleAspLeuLysMetAlaIle 476  
|||||  
101 TTGGAGAGCTGGATTAACCAAGGCATGATCTCATTAATAATGCCATT 150  
477 ProGluLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAs 493  
|||||  
151 CCAGAACTCATGAGGAGGACGTACATTTGTCATGCTGGATCTGGGGA 200  
493 pProIlePheGluGlyTrpMetArgSerThrGluSerSerTyrLysAspL 510  
201 TCCAATTTTGAAGGCTGGATCGAGATCTACCGAGCTAGTTACAGGATA 250  
510 ysPheArgGlyTrpValGlyPheSerValProValSerHisArgIleThr 526  
|||||  
251 AATTCGCGGATGGGTGGATTTACTGCCCAATTTCCACACAATAACT 300  
527 AlaGlyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLe 543  
|||||  
301 GCATGTCGGATATATGTTAATGCCATCCATGTTTGAACCTTGGGTCT 350  
543 uAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValValHisGlyT 560



```

|||||
351 TATCAACTATTATGCTATGCATATGGTACAGTTCCGTGACTTCATGAA 400
|||||
560 hrglygly 562
|||||
401 CTGGGGGC 408

seq_name: gb_est1.AW759569

seq_documentation_block:
LOCUS AW759569 618 bp mRNA 18-JUL-2000
DEFINITION Gm-cl027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
SYNTHASE PRECURSOR ; , mRNA sequence.
ACCESSION AW759569.1 GI:7691442
VERSION AW759569.1
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 618)
Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Willie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Insert Length: 1230 Std Error: 0.00
High quality sequence stop: 411.
Location/Qualifiers
1. 618
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl027-4241"
/collection="Gm-cl027"
/tissue_type="cotyledons of 3- and 7-day-old Williams
seedlings"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; This cDNA library was constructed from mRNA isolated
from cotyledons of 3- and 7-day-old Williams seedlings
which were propagated on paper towels with distilled
water. The cotyledons were flash-frozen in liquid
nitrogen, then lyophilized for 72 hours. Unequal amounts
of mRNA was used for cDNA synthesis. Stratagene's cDNA
Synthesis Kit (catalog number 200401) was used to
synthesize the cDNA. First- stranded synthesis was
performed with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An anchor
nucleotide (V=A, C, or G) was added to the 3' end of the
primer [GAGAGAGAGAGAGAGAACCTAGCTCG(T)18] to anchor
the primer at the 5' end of the poly(A) tract. After
second- strand synthesis, the cDNA ends were filled in
with cloned pfu DNA, ligated to EcoRI adapters and
subsequently electroporated. The xhoI site within the

```







564 ArgAspThr 566  
:::|||||  
606 AAGGATACA 614

seq\_name: gb\_est1:AW599325

seq\_documentation\_block: 592 bp mRNA 22-MAR-2000  
LOCUS AW599325 EST  
DEFINITION gb13e02.y1 Moss EST library PPN Physcomitrella patens cDNA clone  
PEP\_SOURCE\_ID:PPN091503 5' similar to TR:O24398 O24398 STARCH  
SYNTHASE ;, mRNA sequence.

ACCESSION AW599325  
VERSION AW599325.1 GI:7286838  
KEYWORDS EST  
SOURCE Physcomitrella patens.  
ORGANISM Physcomitrella patens.

REFERENCE 1 (bases 1 to 592)  
AUTHORS Quatrano, R., Bashlades, S., Cove, D., Cuming, A., Knight, C., Clifton  
, S., Marra, M., Hillier, L., Pape, D., Martin, J., Wyllie, T., Underwood  
, K., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T.,  
Steptoe, M., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,  
Waterston, R. and Wilson, R.  
Leeds/Wash U Moss EST Project  
Unpublished (1999)  
Contact: Ralph Quatrano  
Leeds/Wash U Moss EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

TITLE  
JOURNAL  
COMMENT

Libraries were constructed by Dr. Stavros Bashlades as part of the  
Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and  
Washington Univ. in St. Louis (USA) DNA sequencing by: Washington  
University Genome Sequencing Center For information on obtaining a  
clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)  
Seq primer: -40RP from Gibco  
High quality sequence stop: 419.

FEATURES  
source

/organism="Physcomitrella patens"  
/db\_xref="taxon:3218"  
/clone="PEP\_SOURCE\_ID:PPN091503"  
/clone\_lib="Moss EST library PPN"  
/tissue\_type="protonemata: 7 day old tissue auxin treated"  
/lab\_host="Dhl0B"  
/note="vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
XhoI; Construction of the cDNA library was carried out  
using Stratagene's 'Unizap - cDNA synthesis kit'. cDNA was  
constructed using an oligo dt primer/linker that contains  
a XhoI site within it. Following ds cDNA synthesis,  
EcoRI adapters were ligated to the blunt ends and sample  
was digested with XhoI. The result is cDNA with an EcoRI  
sticky end on one side and a XhoI sticky end on the other.  
This cDNA was ligated directionally in Unizap arms. The  
vector is designed containing the pBluescript sequence as  
well as lambda DNA and cDNA is cloned within this  
pBluescript sequence. The vector was then packaged using  
Gold gigapackaging extracts. Library was grown in XLBlue  
MRF' cells and amplified. The library was excised by mass  
excision using Stratagene's 'Mass excision kit' that uses  
exassist as a helper phage that releases the pBluescript  
sequence and circularises it as single stranded plasmids  
that are then packaged (by helper phage) and secreted out  
of the host cell as phagemids. SOLR cells were transformed  
with phagemids and the library was plated out on LB-amp  
plates to select for transformants. Approximately 1,000  
,000 colonies were grown and recovered. The double  
stranded plasmid library was recovered by using Qiagen  
Midi prep kit. 2 micro grams of each library were used to

transform Dhl0B cells by electroporation."  
BASE COUNT 142 a 126 c 161 g 160 t 3 others  
ORIGIN

alignment\_scores:  
Quality: 638.50 Length: 199  
Ratio: 3.893 Gaps: 2  
Percent Similarity: 82.412 Percent Identity: 61.307

alignment\_block:

US-09-674-824-2 x AW599325 ..

Align seg 1/1 to: AW599325 from: 1 to: 592

427 TyrSerValAspLeuSerGlyLysAlaLysCysLysAlaGluLeuGl 443  
:::|||||  
11 CACTCTTTAGAC.....ACAGGTAAGGCGCATGCAAGGCTGCTTTACA 54  
443 nLysGluLeuGlyLeuProValArgGluAspValProLeuLeuGlyPheI 460  
55 AAGGGAGTTGGGCTTGGCGGTGGCGGTGACGTCCTCCATTCCTGGATTCA 104  
460 leGlyArgLeuAspTyrGlnLysGlyIleAspLeuLeuLysMetAlaIle 476  
105 TAGGTCCCTTGTATCACCACCAAGGTTATTGACATATAGGCCAAGCTATG 154  
477 ProGluLeuMetArgGluAspValGlnPheValMetLeuGlySerGlyAs 493  
155 CCATGGATGATGGACCAAGATATCCCAACTGGTATGCTGGCCTGGCCG 204  
493 pProIlePheGluGlyTrpMetArgSerThrGluSerSerTyrLysAspL 510  
205 GAAAGATTATTAAGACATGTTGAGGCATTTTGAAGTTCTCATCGAGACA 254  
510 ysPheArgGlyTrpValGlyPheSerValProValSerHisArgIleThr 526  
255 AAGTTCGGGGATGGTTGGTTTCTGTGACCACTTCCCATCGGATTACC 304  
527 AlaGlyCysAspIleLeuLeuMetProSerArgPheGluProCysGlyLe 543  
305 GCAGGTGTTGATATTGTTGATGCCATCTCGGTTTGAACCCCTGTGGCTT 354  
543 uAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProValValHisGlyT 560  
355 GAACCACTGATGCGATGCGGTATGCGACAGTCCAGTAGTCATGCTG 404  
560 hrGlyGlyLeuArgAspThrValGluThrPheAsnPropheGlyAlaLys 576  
405 TTGTGGCTGAAAGATACAGTCAGTCAGTCGTTCAATCCCTTC..... 445  
577 GlyGluGluGlyThrGlyTrpAlaPheSerProLeuThrValAspLysMe 593  
446 AACGAGTCAGGCTCCTGANGATGAGCTNTCGAACACTACATGTGGACGCTT 495  
593 tLeuTrpAlaLeuArgThrAlaMetSerThrPheArgGluHisLysProS 610  
496 TATCATCCCTTGGGATGCAATTTGGACCTACCGTATTCAGGACTA 545  
610 erTrpGluGlyLeuMetLysArgGlyMetThrLysAspHisThrTrp 625  
546 GTTGAAGGCGCATTCAGCAACGATGATGCTCTCANGACTTGAGCTGG 592

seq\_name: gb\_est1:AU075475

seq\_documentation\_block:

LOCUS AU075475 396 bp mRNA EST  
DEFINITION AU075475 Rice cDNA from immature leaf including apical meristem  
ORYZA SATIVA cDNA clone E60759\_1A, mRNA sequence.  
ACCESSION AU075475  
VERSION AU075475.1 GI:5056096  
KEYWORDS .EST.  
SOURCE Oryza sativa.



ORGANISM	Oryza sativa
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
TITLE	1 (bases 1 to 396)
JOURNAL	Sasaki,T. and Yamamoto,K.
COMMENT	Rice cDNA from immature leaf including apical meristem Unpublished (1997) Contact: Takuji Sasaki National Institute of Agrobiological Resources Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan Tel: 81-298-38-7441 Fax: 81-298-38-7468 Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/ PROJECT = 'RGP',
FEATURES	Location/Qualifiers
source	1..396 /organism="Oryza sativa" /strain="Nipponbare" /db_xref="taxon:4530" /clone="E60759.1a" /clone_lib="Rice cDNA from immature leaf including apical meristem" /dev_stage="Immature" /note="Organ: leaf; immature leaf including apical meristem (under long day condition)"
BASE COUNT	104 a 72 c 103 g 116 t
ORIGIN	1 others

```

alignment_scores:
  Quality: 613.00      Length: 132
  Ratio: 4.827         Gaps: 1
  Percent Similarity: 96.212  Percent Identity: 88.636

alignment_block:
US-09-674-824-2 x AU075475  ..

Align seg 1/1 to: AU075475 from: 1 to: 396

440 AlaGluLeuGlnLysGluLeuGlyLeuProValArgGluAspValProLe 456
|||||
2 GCTGAATTCAGAAAGCAGCTGGTTTACCTATAAGGCCCGCATGTGCCCT 51
|||||

456 uileGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLeuLeL 473
|||||
52 GATGTGCTTTATTGGAAGATTGGACTATCAAAAAGGCATTGATCAATTA 101
|||||

473 yMetAlaIleProGluLeuMetArgGluAspValGlnPheValMetLeu 489
|||||
102 AACTTGCCTTCAGATCTCATGCGGGCAATATTCAATTCTGCTGCTT 151
|||||

490 GlySerGlyAspProIlePheGluGlyTyrMetArgSerThrGluSerSe 506
|||||
152 GGATCTGGTGACCCAGGTTTTCAAGGATGGATGAGATCCACAGAAATCAGG 201
|||||

506 rTyrLysAspLysPheArgGlyTyrValGlyPheSerValProValSerH 523
|||||
202 GTACAGGGATAAATTTCGTGGATGGTTGGATTAGTGTCCAGTTTCCC 251
|||||

523 isArgIleThrAlaGlyCysAspIleLeuLeuMetProSerArgPheGlu 539
|||||
252 ACCGATACTGCAGGTGGCGATATATTGTTGATGCCATCCAGAAATCGAA 301
|||||

540 ProCysGlyLeuAsnGlnLeuTyrAlaMetGlnTyrGlyThrValProva 556
|||||
302 CCTTGTGGCCTCAATCAGCTATATCTATGCAATATGTCACAGTGCCTGN 351
|||||

556 lValHisGlyThr..GlyClyLeuArgAspThrValGluThrPhe 570
|||||
352 TGTTCATGGGAATCAGAGGCCCTCAGAGATACAGTGAGAGATTTT 396
|||||

```

```

seq_name: gb_est1:AW438020

seq_documentation_block:
  LOCUS      AW438020          411 bp      mRNA          EST          14-FEB-2000
  DEFINITION  ST83F01 Pine Triplex shoot tip library Pinus taeda cDNA clone
  ST83F01, mRNA sequence.
  ACCESSION  AW438020
  VERSION    AW438020.1 GI:6973326
  KEYWORDS   EST.
  SOURCE     loblolly pine.
  ORGANISM   Pinus taeda
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

REFERENCE
  AUTHORS    Whetten,R.W., Kintaw,C.S., Retzel,E. and Sederoff,R.R.
  TITLE      The Pine Gene Discovery Project
  JOURNAL    Unpublished (1999)
  COMMENT    Contact: Ross Whetten
              Forest Biotechnology Group
              North Carolina State University
              Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh
              ,NC, 27695-8008
              Tel: 919-515-7800
              Fax: 919-515-7801
              Email: rosswhet@unity.ncsu.edu
  Seq primer: 5' lambda Triplex2 Sequencing Primer.

FEATURES             Location/Qualifiers
     source            1..411
                     /organism="Pinus taeda"
                     /db_xref="taxon:3352"
                     /clone="ST83F01"
                     /clone_lib="Pine Triplex shoot tip library"
                     /lab_host="E. coli BM25.8"
                     /note="Organ: shoot tips; Vector: Lambda Triplex; Site_1:
                     Sfil (A); Site_2: Sfil (B); Shoot tips (approx. 2 cm from
                     apex) were collected during the spring, frozen and used
                     for mRNA isolation. The SMART-PCR method (Clontech) was
                     used to prepare a library from 1 ug total RNA, using the
                     Lambda Triplex vector. Plasmid subclones in pTriplex were
                     recovered by cre-lox excision in E. coli strain BM25.8 and
                     sequenced from the 5' end."
     BASE COUNT       111 a      78 c      95 g      113 t      14 others
     ORIGIN

alignment_scores:
  Quality: 555.00      Length: 137
  Ratio: 4.549        Gaps: 0
  Percent Similarity: 89.051      Percent Identity: 72.993

alignment_block:
  US-09-674-824-2 x AW438020      ..

  Align seg 1/1 to: AW438020      from: 1 to: 411

405 GlytIleValAsnGlyIleAspIleAsnAspTrpAsnProThrThrAspLy 421
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1  GGCATCAGTAATGGAATGTGATGTGTGATGGAGTAATCCAGTTTCAGACAA 50

421 sCysLeuProHisHisTyrSerValAspAspLeuSerGlyLysAlaLysC 438
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 ACATGTCCCATTCCTCAGTACTCTGTAGAGGACCITTCGTGGAAGCAAAAT 100

438 ySLysAlaGluLeuGlnInLysGluLeuGlyLeuProValArgGluAspVal 454
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 GCAAGGCAGCTCTACAAATTCGAGCTTGGACTTCTGTTCCAGCATGTG 150

455 ProLeuIleGlyPheIleGlyArgLeuAspTyrGlnLysGlyIleAspLe 471
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 CCGCTGATTGGTTTCATTGGAAGATTGGACTACCAAAAGGACTTGATCT 200

471 uIleLysMetAlaIleProGluLeuMetArgGluAspValGlnPheValM 488
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```







---



**THIS PAGE BLANK (USPTO)**